



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CLASIFICACIÓN Y DISTRIBUCIÓN DE ERRORES DE AUTOMATIZACIÓN EMPLEANDO TÉCNICAS DE MACHINE LEARNING EN LA EMPRESA DIGITAL WARE

AUTORA: DEISY KARINA LOZANO VERA

INGENIERÍA EN TELECOMUNICACIONES
DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA,
SISTEMAS Y TELECOMUNICACIONES
FACULTAD DE INGENIERÍAS Y ARQUITECTURA
UNIVERSIDAD DE PAMPLONA
PAMPLONA, noviembre 25 de 2021



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CLASIFICACIÓN Y DISTRIBUCIÓN DE ERRORES DE AUTOMATIZACIÓN EMPLEANDO TÉCNICAS DE MACHINE LEARNING EN LA EMPRESA DIGITAL WARE

Trabajo de grado presentado como requisito para optar al título de
INGENIERO EN TELECOMUNICACIONES

Directora: KELLY JOHANNA TORRES NAVARRO

Ingeniera en Telecomunicaciones

Especialista en Seguridad Informática en formación

INGENIERÍA EN TELECOMUNICACIONES
DEPARTAMENTO DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA,
SISTEMAS Y TELECOMUNICACIONES
FACULTAD DE INGENIERÍAS Y ARQUITECTURA
UNIVERSIDAD DE PAMPLONA
PAMPLONA, noviembre 25 de 2021



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



PROGRAMA DE INGENIERÍA EN TELECOMUNICACIONES
TRABAJO PRESENTADO PARA OPTAR POR EL TÍTULO DE INGENIERO EN
TELECOMUNICACIONES

TEMA:

**CLASIFICACIÓN Y DISTRIBUCIÓN DE ERRORES DE AUTOMATIZACIÓN EMPLEANDO
TÉCNICAS DE MACHINE LEARNING EN LA EMPRESA DIGITAL WARE**

FECHA DE INICIO DEL TRABAJO:

FECHA DE TERMINACIÓN DEL TRABAJO:

NOMBRES Y FIRMAS DE AUTORIZACIÓN PARA LA SUSTENTACIÓN:

DEISY KARINA LOZANO VERA

AUTOR

KELLY JOHANNA TORRES NAVARRO

DIRECTORA TRABAJO GRADO

EDWIN MAURICIO SEQUEDA ARENAS

DIRECTOR DE PROGRAMA



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



JURADO CALIFICADOR:

GERMÁN ARLEY PORTILLA
ING. ELECTRÓNICO

KELLY JOHANNA TORRES NAVARRO
ING. EN TELECOMUNICACIONES

HERNANDO JOSÉ VELANDIA
ING. ELECTRONICO



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



AGRADECIMIENTOS

Doy gracias a Dios por ser la fuente de todo el empeño realizado en el transcurso de mi formación, por ser quien guiara mis decisiones y mis procesos, a mis padres por todo el sacrificio que han hecho para que yo lograra cumplir uno de los sueños más grandes, por su constante apoyo, ánimos en momentos difíciles y ser la fuente de inspiración para superarme cada día como persona y profesional. A mi hermana, por ser el mejor ejemplo que pude tomar, con sus actitudes al enfrentar inconvenientes me ha enseñado a siempre perseguir mis sueños a pesar de las adversidades. Quiero agradecer a mi mano derecha en todo este proceso, mi novio, que siempre ha estado ahí apoyándome y brindándome todo su amor incondicional para llegar donde estoy. Y a mis abuelos, que a pesar de que no pudieron verme cumplir este sueño, sé que estarán muy orgullosos de lo que he logrado y esto también va dedicado especialmente a ellos por lo importante que fueron en mi formación personal y académica. Mil gracias por todo.

Y por supuesto debo agradecer, a mis formadores, quienes con sus conocimientos fueron ayudándome a formar mi personalidad y mi parte académica, siendo arte fundamental de todo este lindo proceso, también a mis compañeros que siempre estuvimos unidos apoyándonos para así lograr este gran propósito de ser llamados ingenieros en telecomunicaciones.

Y finalmente agradecer a Digital Ware y a todo su equipo por permitirme crecer profesionalmente brindándome conocimientos para seguir adquiriendo experiencia y sobre todo crecer como persona.

Muchísimas gracias.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Tabla de contenido

RESUMEN	12
ABSTRACT	13
CAPITULO I. INTRODUCCION	14
PROBLEMA	15
1.1 JUSTIFICACION	15
1.2 DIAGRAMA CAUSA-EFECTO	16
1.3.1 OBJETIVOS	17
1.3.1.1 OBJETIVO GENERAL	17
1.3.1.2 OBJETIVOS ESPECÍFICOS	17
CAPITULO II. MARCO INSTITUCIONAL	18
2.1 DIGITAL WARE	18
2.1.1 MISIÓN	18
2.1.2 VISIÓN	18
2.1.3 CARGO DESEMPEÑADO	19
CAPITULO III. MARCO REFERENCIAL	23
3.1 ESTADO DEL ARTE	23
CAPITULO IV. MARCO TEORICO	26
1. AUTOMATIZACION	26
2. AUTOMATIZACION DE PRUEBAS	26
2.1 Tipos de pruebas automatizadas	26
2.1.1 Unitarias	26
2.1.2 Funcionales	27
2.1.3 Pruebas de Integración	27



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



2.1.4	Pruebas de punta a punta	27
2.1.5	Pruebas de Regresión	28
2.1.6	Pruebas de humo	28
2.1.7	Pruebas de aceptación	28
2.1.8	Pruebas de rendimiento	28
3.	SELENIUM	29
4.	BASES DE DATOS SQL Y ORACLE	29
4.1	Oracle	29
4.2	SQL	29
5.	LENGUAJE C#	30
6.	NET FRAMEWORK	31
7.	NET CORE	33
8.	INTELIGENCIA ARTIFICIAL	34
8.1	Técnicas de inteligencia artificial	35
9.	MACHINE LEARNING	37
9.1	Tipos de Machine Learning	38
9.1.1	Aprendizaje Supervisado	39
9.1.2	Aprendizaje no supervisado:	51
9.1.3	Aprendizaje por refuerzo:	55
10.	ML.NET (MACHINE LEARNING .NET)	56
10.1	Características	56
10.2	Componentes de ML.NET	57
CAPITULO V. METODOLOGIA.....		60
5.1	DESCRIPCION DEL LUGAR DE TRABAJO.	60
5.2	DESARROLLO DE LA METODOLOGIA.....	62
FASE 1.....		62



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



FASE 2.....	66
FASE 3.....	86
FASE 4.....	104
CAPITULO VI. RESULTADOS	111
CAPITULO VII. CONCLUSIONES.....	119
REFERENCIAS BIBLIOGRAFICAS.....	121



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



ILUSTRACIONES

Ilustración 1. Planteamiento del problema [Propia]	16
Ilustración 2. Que permite hacer SQL [18]	30
Ilustración 3. C# [22]	31
Ilustración 4. Arquitectura NET Framework [24]	32
Ilustración 5. Estructura NET CORE [27]	34
Ilustración 6. Inteligencia Artificial [31]	36
Ilustración 7. Comportamiento Machine Learning [34]	38
Ilustración 8. Tipos de Machine Learning [34]	38
Ilustración 9. Clasificación [36]	39
Ilustración 10. Comportamiento matemático de clasificación binaria [58]	40
Ilustración 11. Clasificación Binaria [37]	42
Ilustración 12. Comportamiento clasificación multiclase [59]	43
Ilustración 13. Clasificación multiclase [38]	43
Ilustración 14. Función Logística [39]	44
Ilustración 15. Modelo regresión lineal [39]	45
Ilustración 16. Maximizar margen [60]	46
Ilustración 17. Truco Kernel [40]	47
Ilustración 18. Separación Hiperplanos [60]	47
Ilustración 19. Árbol de decisión [60]	48
Ilustración 20. Comportamiento de Regresión Lineal [61]	49
Ilustración 21. Regresión no lineal [61]	51
Ilustración 22. Aprendizaje no supervisado [45]	52
Ilustración 23. Clustering [46]	53
Ilustración 24. Técnica KMEANS [47]	53
Ilustración 25. Reducción de dimensionalidad [48]	54
Ilustración 26. Técnicas de Reducción de dimensionalidad [49]	55
Ilustración 27. Aprendizaje por refuerzo [50]	55
Ilustración 28. Metodología [Propia]	61
Ilustración 29. Pruebas automatizadas [Propia]	66
Ilustración 30. Ejemplo código pruebas automatizadas [Propia]	67
Ilustración 31. Datos almacenados en MTM [Propia]	68
Ilustración 32. Caso en MTM [Propia]	68
Ilustración 33. Creación Pipeline [Propia]	69
Ilustración 34. Grupo Agentes [Propia]	70
Ilustración 35. Canalización [Propia]	70
Ilustración 36. Parámetros pipeline [Propia]	71
Ilustración 37. Compilar Solución [Propia]	72



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



Ilustración 38. Actualizar Tabla [Propia]	73
Ilustración 39. TestAgent [Propia]	74
Ilustración 40. Pruebas a lanzar [Propia]	75
Ilustración 41. Actualizar Tabla 2 [Propia]	76
Ilustración 42. Test Agent 2 [Propia]	77
Ilustración 43. Conjunto Pruebas 2 [Propia]	78
Ilustración 44. Poner en cola [Propia]	79
Ilustración 45. Compilaciones del Pipeline [Propia]	79
Ilustración 46. Pipelines [Propia]	80
Ilustración 47. Proceso compilación desarrollado [Propia]	81
Ilustración 48. Resultados Pruebas [Propia]	82
Ilustración 49. Errores Presentados [Propia]	83
Ilustración 50. Errores obtenidos de los resultados pruebas automatizadas [Propia]	84
Ilustración 51. Datos adecuados formato txt con tabulaciones [Propia]	85
Ilustración 52. Visual Studio [Propia]	86
Ilustración 53. Aplicación de consola [Propia]	87
Ilustración 54. Nombre Proyecto y Solución [Propia]	87
Ilustración 55. Versión Net. Core [Propia]	88
Ilustración 56. Ambiente de trabajo [Propia]	88
Ilustración 57. Solución Proyecto [Propia]	89
Ilustración 58. Administrador de NuGets [Propia]	89
Ilustración 59. NuGet ML [Propia]	90
Ilustración 60. Instalación NuGet ML [Propia]	90
Ilustración 61. Abrir carpeta en Explorador de archivos [Propia]	91
Ilustración 62. Directorio Data [Propia]	92
Ilustración 63. Copia data adecuada en txt. [Propia]	92
Ilustración 64. Aplicación en explorador de soluciones [Propia]	92
Ilustración 65. Agregar nuevo elemento [Propia]	93
Ilustración 66. Clase MistakeData [Propio]	94
Ilustración 67. Librerías clase MistakeData [Propia]	94
Ilustración 68. Código Clase MistakeData [Propia]	95
Ilustración 69. Clase MistakePrediction [Propia]	96
Ilustración 70. Librerías clase Program [Propia]	97
Ilustración 71. Contexto ML.NET [Propia]	97
Ilustración 72. Ubicación archivo txt [Propia]	98
Ilustración 73. LoadData [Propia]	99
Ilustración 74. Model [Propia]	100
Ilustración 75. Evaluate [Propio]	101
Ilustración 76. Predicción [Propia]	103



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Ilustración 77. Main [Propia]	104
Ilustración 78. Error validación #1 [Propia]	105
Ilustración 79. Predicción algoritmo error #1 [Propia]	105
Ilustración 80. Error validación #2 [Propia]	106
Ilustración 81. Predicción error #2 [Propia]	106
Ilustración 82. Error validación #3 [Propia]	107
Ilustración 83. Predicción error #3 [Propia]	107
Ilustración 84. Error validación #4 [Propia]	108
Ilustración 85. Predicción error #4 [Propia]	108
Ilustración 86. Error validación #5 [Propia]	109
Ilustración 87. Predicción error #5 [Propia]	109
Ilustración 88. Evaluación del Modelo [Propia]	110
Ilustración 89. Resultado Contexto [Propia]	112
Ilustración 90. Ruta data [Propia]	112
Ilustración 91. Data composición [Propia]	113
Ilustración 92. Fracción Data [Propia]	113
Ilustración 93. Validación carga y entreno correcto del modelo [Propia]	113
Ilustración 94. Consumo CPU [Propia]	114
Ilustración 95. Memoria vs Tiempo [Propia]	114
Ilustración 96. Consumo de recursos	115
Ilustración 97. Visualización de la predicción [Propia]	116
Ilustración 98. Precisión sistema [Propia]	117
Ilustración 99. Area bajo la curva [Propia]	117
Ilustración 100. Puntaje F1 [Propia]	118

TABLAS

Tabla 1. Regla de 3A's [Propia].....	27
Tabla 2. Paso a paso para crear un modelo con ML, NET [54].....	57
Tabla 3. Resultados en métricas [Propia].....	117



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



RESUMEN

En el presente trabajo se expone el desarrollo de un sistema de clasificación y distribución de errores por medio de técnicas de Machine Learning la cual permita dar solución a una de las problemáticas más comunes en el área de automatización de la empresa Digital Ware como lo es la dificultad en la revisión y reajuste de pruebas automatizadas debido al alto manejo de pruebas e información aumentando el tiempo requerido haciendo lento el proceso. Para ello se plantea que este sistema haga uso de la selección y distribución de errores teniendo como referencia los resultados obtenidos en lanzamientos masivos en automatización a través de pipeline de Azure conectados al proyecto empresarial, con los algoritmos y técnicas facilitadas en las librerías de .NET tales como ML.NET y aplicaciones de consola como NET CORE; donde el algoritmo tenga la entrada de datos previamente adecuada y estructurada, haciendo uso de los formatos de trata de datos como .txt separando columnas con tabulaciones para fácil lectura por parte de las librerías, facilitando su extracción, y clasificándolo a través del modelo binario con técnica que mas estable como regresión logística, entre otras. Finalmente se obtiene el sistema de clasificacion funcional y evaluado por las métricas dadas estrictamente para clasificaciones binarias dando como resultado un alto porcentaje de precisión y confidencialidad.

Palabras clave: Machine Learning, pruebas automatizadas, ML.NET, Clasificación, NET.CORE



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



ABSTRACT

In this paper, the development of a system for classifying and distributing errors through Machine Learning techniques is exposed, which allows solving one of the most common problems in the automation area of the Digital Ware company, such as Difficulty in the revision and readjustment of automated tests due to the high handling of tests and information increasing the time required making the process slow. For this, it is proposed that this system make use of the selection and distribution of errors, taking as a reference the results obtained in massive automation launches through Azure pipeline connected to the business project, with the algorithms and techniques provided in the .NET libraries such as ML.NET and console applications such as NET CORE; where the algorithm has previously adequate and structured data entry, making use of data processing formats such as .txt, separating columns with tabulations for easy reading by libraries, facilitating its extraction, and classifying it through the binary model with more stable technique such as logistic regression, among others. Finally, the functional classification system is obtained and evaluated by the metrics strictly given for binary classifications, resulting in a high percentage of precision and confidentiality.

Keywords: Machine Learning, automated tests, ML.NET, Classification, NET.CORE



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CAPITULO I. INTRODUCCION

Actualmente en el desarrollo de software y realización de pruebas automatizadas está en constante crecimiento gracias a su eficacia en evaluar la calidad del producto, de forma paralela a su crecimiento, se empiezan a manejar grandes cantidades de datos y creación de pruebas automatizadas de alta demanda, generando grandes cargas a los programadores encargados de la creación, corrección y soporte constante a dichas pruebas.

Esta última acción depende mucho de la revisión unificada de cada una de las pruebas lanzadas para notificar cuales son las que necesitan correcciones de errores; el realizar esta revisión exhaustiva se requiere demasiado tiempo y en conjunto a la gran cantidad de pruebas que se manejan da como resultado la problemática más destacada en los equipos de automatización.

Actualmente una solución que dan los avances tecnológicos, es hacer uso de las máquinas con previo entrenamiento que sean capaces de tomar decisiones que simplifiquen las intervenciones de los humanos, limitándolos solo a extraer los datos y de ahí se parte ante la toma de decisiones, esta tecnología es conocida como Machine Learning, la cual presenta varios modelos de los cuales se elige el adecuado para la funcionalidad del sistema, y basándose en ese modelo se presentan infinidad de algoritmos que con comportamientos matemáticos diferentes permiten acoplarse a lo que se requiere. Por ello es importante saber al comienzo del uso de esta tecnología cual es el objetivo principal y funcional del proyecto donde se aplicará Machine Learning para así acoplarlo a los mejores modelos y técnicas más precisas para tener un modelo con alto porcentaje en efectividad.

Es por ello que para este proyecto se plantea el desarrollo de un sistema de clasificación de errores con Machine Learning haciendo uso de herramientas como ML.NET y NET.CORE, que permita rediseñar que tipo de error se presenta en las pruebas lanzadas y se tenga acceso instantáneo con previo conocimiento para aplicar el soporte, disminuyendo así la carga de los programadores y el tiempo necesario para realizar soporte.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



PROBLEMA

1.1 JUSTIFICACIÓN

Hoy en día las empresas tienen como principal objetivo dar productos y servicios de alta calidad que a través del tiempo con sus constantes cambios tecnológicos deriva a la implementación de actualizaciones mejorando la experiencia del cliente, por lo cual para que dichas experiencias sean positivas internamente en la empresa se deben hacer pruebas de calidad funcionales con todos los parámetros dando el aval de que se liberen cada uno de esos paquetes de actualizaciones, es ahí donde el testing es de suma importancia ya que a través de esta técnica se puede automatizar las pruebas de calidad de los productos y servicios.

Debido a la alta demanda que requiere el uso esta técnica donde se desarrollan contantes pruebas programadas por código donde las acciones permiten evaluar el funcionamiento de los productos, en el área de automatización se manejan dos tipos de pruebas, con pruebas automatizadas dirigidas hacia programas del software empresarial Ophelia y las pruebas web que son dirigidas hacia las páginas web empresariales. [1]

Con el manejo de estos dos tipos, es constante observar muchos lanzamientos con gran número de pruebas, da como resultado demasiados datos que analizar para hacer el reajuste correspondiente, dificultando el acceso a las pruebas que obtuvieron errores y que tipo de error presenta, es por ello que al revisar una por una se toma demasiado tiempo y hace difícil la rápida respuesta por parte del equipo de automatización ante las asignaciones por parte de superiores.

Como respuesta a la problemática presentada, a través de este proyecto se desarrollará un sistema como base principal el uso de Machine Learning haciendo uso de tecnologías .NET como ML.NET y NET. CORE, en el cual se realiza la elección de los errores resultantes en pruebas automatizadas lanzadas en las plataformas de la empresa y páginas web, esto con el fin de clasificarlos de acuerdo a ciertas características específicas que permitan detectarlos con un alto porcentaje de efectividad.

Para ello, se hace uso de datos con ciertas características y etiquetas que permiten la clasificación de errores que se presenten por fallas del sistema o errores programados por nosotros, facilitando así el acceso a cada uno de ellos, realizar el ajuste correspondiente y agilizar la revisión por parte del equipo de automatización. [1]

Para la realización de este proyecto, se propone un paso a paso conformado de cuatro etapas, donde la primera se basará en la adquisición de herramientas necesarias para la codificación del algoritmo, en las que se destacan Visual Studio y sus respectivas extensiones, bases de datos,



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



librerías ML.NET, aplicaciones de consola como Net. Core, ML. Data entre otras.

La segunda etapa se basa en la recolección de data necesaria efectuando el lanzamiento de pruebas automatizadas tantas que sean necesarias para obtener efectividad de predicción alta y extraer los resultados erróneos obtenidos.

En la tercera etapa toda la información obtenida será configurada para que el algoritmo responda correctamente ante la presencia de errores, de tal manera que sea ingresada en un formato que la maquina pueda entender cuál es la base para clasificar los errores.

Y finalmente se da la etapa de aplicación donde se montará las partes implicadas en el proceso, dando al final el programa estructurado completamente, donde se unan todas las partes presentadas como entrenamiento, predicción e ingreso de datos, de donde se obtiene los datos de entrada y cuál es la predicción que clasifique el error.

1.2 DIAGRAMA CAUSA-EFECTO

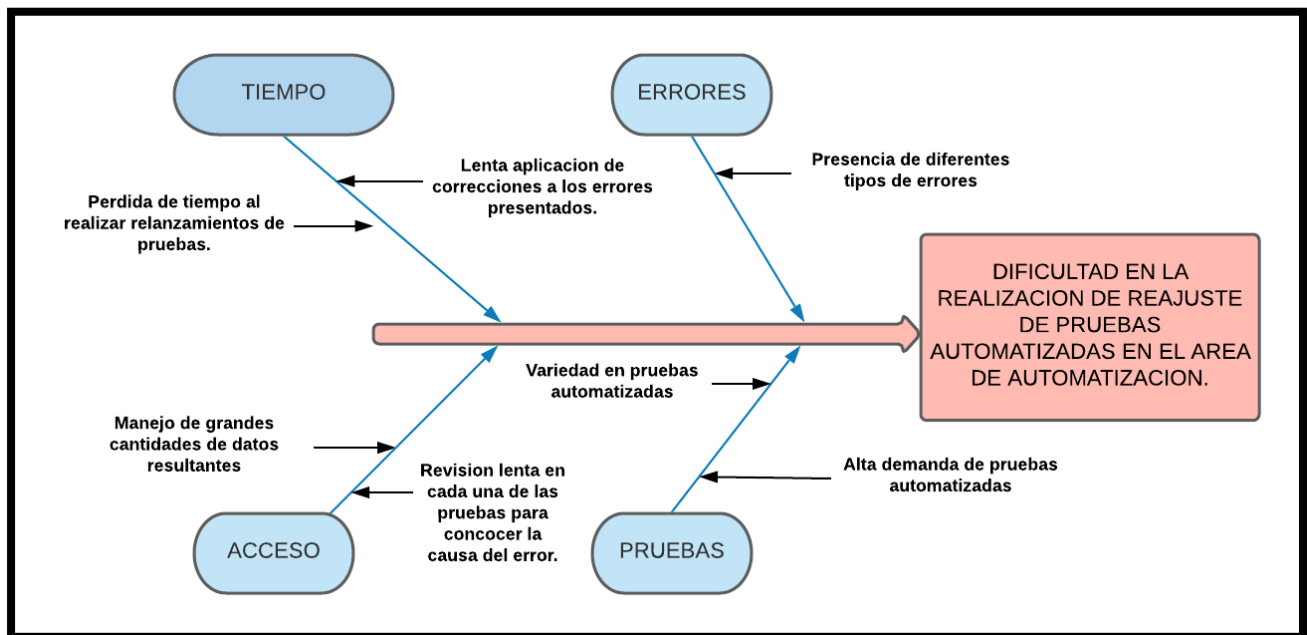


Ilustración 1. Planteamiento del problema [Propia]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



1.3 DELIMITACION

1.3.1 OBJETIVOS

1.3.1.1 OBJETIVO GENERAL

Desarrollar un sistema de clasificación haciendo uso de Machine Learning para distribuir los errores de pruebas automatizadas.

1.3.1.2 OBJETIVOS ESPECÍFICOS

- Obtener conocimientos sobre el manejo de Machine Learning y pruebas automatizadas con las herramientas necesarias para su correcto desarrollo.
- Recopilar la data necesaria para el entrenamiento del algoritmo ante la detección de errores.
- Crear el sistema de clasificación y distribución de errores basándose en la técnica escogida y la data previamente extraída.
- Validar el funcionamiento del sistema desarrollado analizando los resultados obtenidos.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



CAPITULO II. MARCO INSTITUCIONAL

El desarrollo de la práctica empresarial en la empresa DIGITAL WARE fue desarrollada por un tiempo de 6 meses a partir del 6 de julio 2021 hasta el 6 de enero del año 2022 de manera remota en la ciudad de Pamplona, Norte de Santander.

2.1 DIGITAL WARE

Digital Ware es una de las empresas más importantes a nivel nacional e internacional en el ámbito de software de nómina y gestión humana, software para registros en clínicas e ips ya que cuenta con casi tres décadas en el mercado creando sistemas financieros, de nómina, de importación, de salud, de seguridad entre otros. Es catalogada como la casa de software colombiana debido a que facilita la transformación digital en más de 1.000 empresas del sector público y privado en Colombia y Latinoamérica proporcionando software de nómina, gestión humana, ERP, BPM y soluciones especializadas por industrias.

Jorge Camilo Bernal fundo la empresa en el año 1992 y desde ese momento ha recibido más de 11 reconocimientos nacionales e internacionales por ser una de las multinacionales más innovadoras en el sector económico del país aumentando así la economía nacional.[1]

2.1.1 MISIÓN

Mejorar la vida de las personas, y ayudar a las empresas a cumplir sus metas, utilizando el conocimiento, el talento y la tecnología, soportados en visión, pasión, conciencia, disciplina e intensidad. [2]

2.1.2 VISIÓN

Nos convertiremos en la más innovadora, motivadora y servicial empresa de tecnología en América Latina en el desarrollo de plataformas informáticas enfocadas en ERP (Planeación de Recursos Empresariales), HR (Recursos Humanos), HIS (Sistemas de Información Hospitalaria), BPM (Gestión por Procesos Inteligente). También queremos que para el 2023, tener operaciones en 5 países de América Latina, con ventas por US\$ 100 millones. En el 2025, Empresa Pública con operación en toda América y ventas por US\$ 1 Billón. [2]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



2.1.3 CARGO DESEMPEÑADO

El cargo desempeñado en la práctica empresarial corresponde a programar y parametrizar casos de uso asignados por parte del área de calidad correspondientes al proyecto Ophelia y proyecto Smart People, donde se realiza la programación de pruebas automatizadas con unas parametrizaciones previas para evaluar la calidad de las actualizaciones aplicadas a los productos empresariales en KACTUS-SCM y SMART PEOPLE.

Inicialmente se desarrolla el proceso de adaptación correspondiente a la programación de las pruebas automatizadas para testear las funciones principales del programa híbrido Ophelia tales como editar, borrar, añadir y consultar registros que a su vez permita evaluar otras funcionales tales como exportar archivos en Excel, pdf e informes preliminares, estas programaciones en el proyecto son conocidas como CRUD. Para el proyecto correspondiente a Smart People las pruebas automatizadas se realizan para evaluar el estado de los elementos de la página web y los campos donde se ingresan los datos, también se evalúa el comportamiento de las solicitudes tanto en registro como en aprobaciones, lo cual es conocido como flujo.

El lenguaje de programación usado para el desarrollo de las pruebas en ambos proyectos corresponde a C# y trabaja en conjunto con Selenium y Appium, además de ser controlado haciendo uso de Azure y almacenando la información en Test Manager y bases de datos como SQL y Oracle, en cuanto a la herramienta donde se programa, se usan dos versiones de Visual Studio, para Ophelia corresponde a Versión 2019 y para el proyecto web la versión 2017.

Con todo lo anterior, se inició con la creación de nuevas pruebas automatizadas debido a lanzamiento de actualizaciones en el proyecto de Ophelia, se realizó la programación de aproximadamente 50 por semana durante un periodo de mes y medio, dando como resultado 300 pruebas realizadas, donde se dio uso de nuevas herramientas como validación del correcto funcionamiento de lo programado, uso de tecnología que permitió el lanzamiento de varias pruebas en un buen lapso de tiempo. De manera paralela se realizó el proyecto de Microsoft Edge en el cual se realizó la adecuación del Modo internet Explorer en este navegador para el correcto funcionamiento del software empresarial debido a la caducidad que presenta Internet Explorer en la actualidad, para ello se dio uso de políticas de Windows donde se habilitaba la integración de modo internet Explorer y aplicándolo a todo el dominio de la empresa para que todos los colaboradores, de esa manera solo quedaba configurar el navegador para que al ingresar a las páginas que manejan el software de la empresa y este abriera sin ningún problema en la funcionalidad que tenían en el modo internet Explorer, dando un resultado positivo en las ejecuciones, actualmente con las actualizaciones que ha tenido el navegador Edge con el modo internet Explorer que pueda permitir una configuración más dinámica en el futuro que se está empezando a desarrollar.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



Con los numerosos cambios a los que se vio expuesto el área de automatización se me realizó la asignación del reajuste y estabilización del proyecto web Smart People el cual no tenía un mantenimiento correcto en muchos meses en los cuales a nivel de página y HTML se habían realizado muchos cambios, este tipo de pruebas se desarrollan haciendo uso de lenguaje C# , librerías de Appium y Selenium, y algunas funciones de Windows tales como TAB, ENTER, DOWN Y UP, entre otras, y para el lanzamiento masivo se realizan dos configuraciones a nivel de Pipelines como lo es el compilador correspondiente al lanzamiento de los sitios web a usar para la realización de las pruebas, se realizaron dos correspondientes a la base de datos de SQL y otro para la base de datos de Oracle. Al tener las configuraciones de estos sitios web listos, se inició con el lanzamiento masivo de manera local, es decir, en mi maquina únicamente, de pruebas NTC, las cuales corresponden a observar el correcto diseño de la página y de la funcionalidad de los campos de ingreso de texto, check, selección, etc. Para ello, la programación se basa en hacer uso de xpath ya sean relativos o absolutos los cuales ayudan a identificar las características de donde se da el click para validar que el nombre este correcto o donde enviar la función Tab para verificar que el campo de texto o selección estén activos, en este proceso se realizó la reprogramación de aproximadamente 300 pruebas web correspondientes a todos los programas posibles para ingresar desde el aplicativo web. También se realizó una segunda parte de las pruebas web llamadas continuas ya que se programan para validar que los registros, solicitudes y aprobaciones correspondientes a un flujo se desarrolle correctamente, para ello se usaba en la programación variables almacenadas en Test Manager con la información que el robot ingresaba a la página al lanzar la prueba y observar su comportamiento, no obstante, este tipo de pruebas, alrededor de 140, necesitaban una parametrización previa donde se debían cumplir ciertas secuencias de bases de datos que fueron creadas de mi parte basándome en el programa de Kactus donde se almacenaban todos los registros, cumpliendo esa etapa inicial de parametrización de dio la correcta estabilización de todo el proyecto web y con la generación de los reportes para ser informados al jefe del proceso.

También como mejora contante del proyecto se realizó la creación de pipeline en la herramienta Azure para cada uno de los módulos de pruebas previamente estabilizados, en el cual se realizaban 3 configuraciones importantes como lo son: creación del agente en la maquina asignada en las instalaciones de la empresa, la instalación de archivos en disco local correspondientes para el encolamiento de las pruebas en la tabla almacenada en la base de datos de SQL de manera automatizada con su respectiva configuración en el pipeline con la ruta especifica de estos archivos y finalmente se selecciona que módulos se van a lanzar los cuales se almacenan en el Test Manager. Con todas estas configuraciones se realiza la programación de que días se lanzan estas pruebas y a que horas para estar pendiente del proceso de estabilización y si presentan errores nuevamente poder saberlo inmediatamente y realizar el respectivo soporte. Después de realizado este proyecto web, almacenado los informes y



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



estadísticas en porcentaje de pruebas correctas, se envió la información al ingeniero para su revisión y a su vez informado a el área de calidad.

Seguido, se me asigno la tarea de actualizar funciones presentes en el código de tal manera que el 100% del mismo sea configurado usando Selenium y Appium, y a si vez sea configurado en modo headless lo cual permite la ejecución de las pruebas web en segundo plano permitiendo el uso del computador en otras funciones sin ningún problema, para eso se inició con la creación de Una función global en la cual se crearon dos métodos por los cuales se puede iniciar la prueba el primero es en el modo headless y el segundo modo normal, para el primero se realizó la configuración la función de Screenshots para que está fuera realizada basándose únicamente en el navegador para esto se usó la librería de Selenium con su respectiva configuración de gpu, tamaño de la ventana, etc. Al tener este modo sin cabeza configurado y totalmente funcional, se prosigue al cambio de las funciones del proyecto que usaban elementos de Windows por funciones Selenium para que trabajaran al 100% con el navegador.

Como primer paso se realizó el remplazo de la función de Screenshots anterior, por la nueva manejada en Selenium haciendo uso de expresiones regulares para así hacerlo en todo el proyecto sin necesidad de ir por cada módulo presente. Después de ello se realizó la investigación sobre el envío de secuencias del teclado como Tab, enter, Down , up entre otros haciendo uso de Selenium donde se hace uso del xpath del elemento donde se realizara el envío de la secuencia del teclado, después de tener claro la funcionalidad se procede a la realización de pruebas para ver la total funcionalidad de lo implementado para luego hacer su remplazo global en todos los módulos presentes, en cada una de estas secuencias se creó una función global llamada Scroll, Tab, Enter, que son las que más se usan en pruebas automatizadas web. Todo este proceso fue mostrado y reportado al Ingeniero con visto bueno a seguir la mejora del proyecto y dejar una base estructurada de cómo debe ser programada cada una de las pruebas para futuros responsables del proyecto.

De manera diaria, también se realizaban una de las actividades más importantes a nivel empresarial que corresponde a una reunión de aproximadamente una hora con el equipo de automatización donde se tratan los temas más importantes asignados tanto para automatización como para mantenimiento de las bases de datos y servidores, en estas reuniones se tienen tres aspectos importantes tratados como lo es: que se hizo el día de ayer, qué problemas se tuvieron y que se va a realizar el día de hoy; esto con el fin de llevar un seguimiento de las actividades que se realizan y la cantidad de tiempo que debe cumplir las funciones que nos asignan también es un espacio para realizar solicitudes o ser asignados nuevos temas, en el equipo de trabajo esta reunión se hace de 8 de la mañana a 9 de la mañana.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Todo lo anterior abarca las actividades asignadas y realizadas de manera global en la empresa Digital Ware en el equipo de automatización, por lo tanto en el mismo transcurso de las practicas se dio un espacio para la investigación y desarrollo del proyecto de clasificación de errores de las pruebas automatizadas que se venían presentando a medida de la creación de módulos, que permitiera clasificar y distribuir los errores de tal manera que al revisar sea un error de sistema o uno generado por nosotros como programadores, haciendo uso de tecnología Machine Learning de Visual misma herramienta donde se configuran las pruebas conocida como ML.NET ayudando al rápido y eficiente proceso de reajuste.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CAPITULO III. MARCO REFERENCIAL

3.1 ESTADO DEL ARTE

La utilización de la tecnología Machine Learning ha generado innovación tecnológica de alta demanda generando miles de proyectos con distintos propósitos para la mejora de ámbitos laborales generando una alta expectativa en sus implementaciones.

Es por ello, que para este proyecto se tuvieron en cuenta dos estudios internacionales y dos nacionales que sirven para la guía de la implementación de nuestro proyecto de clasificación de errores usando tecnología Machine Learning, estos son:

Autor: Fabian Farola, Carlos Iván Chesñevar, Carlos Alvez, Graciela Etchart, Ernesto, Miranda Silvia Ruiz, Juan José Aguirre, Juan Carlos Teze.

Título: FRAMEWORK SDF MACHINE LEARNING EN TRANSACCIONES FINANCIERAS Y DETECCIÓN TEMPRANA DE FRAUDES

Institución: Universidad Nacional del Sur, Argentina.

Año: 2017

Resumen

En el estudio planteado, se implementa un modelo de detección de fraudes haciendo uso de tarjetas de crédito, implementando modelos predictivos combinando dos importantes técnicas de Machine Learning como lo son el aprendizaje supervisado y no supervisado, que es ingresado a un base de datos, procesada según sus atributos y finalmente implementada para su visualización.

Objetivo

El objetivo es identificar las mejoras aplicables a la clasificación de datos en la clase minoritaria, en entornos de transacciones financieras, para la minimización de falsos positivos. [55]

Gracias a Fabian Farola, Carlos Iván Chesñevar, Carlos Alvez, Graciela Etchart, Ernesto, Miranda Silvia Ruiz, Juan José Aguirre, Juan Carlos Teze, esta tesis permitió conocer las diferentes formas en la que se puede



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



clasificar datos haciendo uso de Machine Learning, en este caso uso dos modelos importantes como lo son aprendizaje supervisado y no supervisado dando una comparativa técnica de los tipos de machine Learning existentes, determinando así cual era la mejor para el proyecto a desarrollar y para determinar la cantidad de casos positivos. en esta tesis explican los modelos y técnicas a usar de acuerdo a la selección que se haga desde regresión logística hasta aprendizaje sin etiquetas, será de gran ayuda para entender mejor lo que ofrece Machine Learning, dando apoyo total a la elección de aprendizaje automático adecuado.

Autor: *Edwar Fabian León Muñoz*

Título: *MODELO DE MACHINE LEARNING PARA LA CLASIFICACIÓN DE ESTUDIANTES DE ACUERDO A SU RENDIMIENTO ACADÉMICO EN EL CENTRO DE IDIOMAS DE LA UNIVERSIDAD NACIONAL DEL SANTA*

Institución: *Universidad nacional de Santa, Perú*

Año: *2020*

Resumen

En el estudio planteado, se implementa un modelo basándose en Microsoft Azure Machine Learning donde se ingresan los datos previamente estructurados de acuerdo a nombres, apellidos, edad, curso, promedio y año, donde se agrupan en una sola data. Se entrena el modelo y verifica la efectividad resultante, también se añade un componente web y un componente API. Al finalizar la estructuración, finalmente se aplica al sistema académico obteniendo un porcentaje de satisfacción por parte de docentes y estudiantes por encima del 70 %.

Objetivo

Según Edwar Fabian León Muñoz, el objetivo de dicha investigación es permitir al Centro de Idiomas de la Universidad Nacional del Santa, realizar la clasificación de los estudiantes según su rendimiento académico, además de contar con una plataforma tecnológica y tener una mejor toma de decisiones. [56]

Con el estudio realizado por Edwar Fabian León Muñoz, permitió profundizar más en la funcionalidad que tiene uno de los softwares más importantes en Machine Learning en Microsoft como lo es Azure, dando al proyecto conocimiento valioso en el manejo de la herramienta y como puede ser aplicada al proyecto, además de dar una estructura correcta de que componentes debe tener la data y el orden para el ingreso y ser usado para entrenar el algoritmo, dando un aporte para la aplicación de nuevas tecnologías de Microsoft en el Machine Learning acorde a la compatibilidad de recursos que se tiene con la empresa.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Autor: Arles Felipe García Maya

Título: Estudio comparativo de técnicas de clasificación binaria con múltiples anotadores

Institución: Universidad Nacional de Colombia. Medellín.

Año: 2020

Resumen

Arles Felipe García Maya, realizó la selección, implementación, pruebas y análisis de las métricas precisión, Recall, F1 Score y ROC de cinco técnicas de aprendizaje de máquina con múltiples anotadores, lo cual permite a la comunidad científica o profesional tener más criterios de selección a la hora de seleccionar alguna de las técnicas mencionadas aquí.

Objetivo

El objetivo es conocer el comportamiento de estas técnicas frente a diferentes bases de datos basados en las métricas de rendimiento obtenidas sobre las pruebas en las diferentes bases de datos evidencian grandes diferencias entre las técnicas para la misma base de datos. [57]

Esta tesis desarrollada por Arles Felipe García Maya, informo de manera detallada las diferentes métricas que existen para evaluar y medir el estado del modelo de machine Learning, permitiéndome escoger de manera mas segura cual de todas es la que necesito y mirar su análisis matemático y como es su comportamiento, entre las que se destacan precision, F1, AUC o area bajo la curva, Recall y ROC, las cuales serán analizadas a lo largo de este proyecto.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CAPITULO IV. MARCO TEORICO

Para poder contextualizar el desarrollo del proyecto es necesario comprender las partes implicadas, conocer su significado y clasificación que permitan entender de forma más clara el desarrollo de la metodología del proyecto, para ello se realiza la investigación de las definiciones de los elementos más importantes acompañado de las características más sobresalientes e ilustraciones que expliquen de manera sencilla el concepto.

1. AUTOMATIZACIÓN

Es una disciplina de control realizada a sistemas para poder controlar los diversos procesos haciendo uso de elementos tecnológicos que evalúan los comportamientos permitiendo el incremento de la producción sin necesidad de la intervención significativa del ser humano.[3]

Usa sistemas de software para poder crear procesos repetitivos con el objetivo de reducir el tiempo en producción obteniendo una optimización de entornos totalmente funcional, donde solo se tengas procesos codificados y no manuales.[4]

2. AUTOMATIZACIÓN DE PRUEBAS

Las pruebas automatizadas son ejecuciones de una serie de comandos en cierto lenguaje de programación que permitan la ejecución de manera automática sin intervención del usuario en todo su proceso. Este tipo de pruebas mejoran la confiabilidad, también aceleran el proceso de prueba, mejoran la cobertura de prueba y brindan confianza a los evaluadores para una prueba de mejor calidad del producto, siendo de las mejores actualmente para probar el funcionamiento de los productos donde se apliquen, como web o programas.[5]

Este tipo de pruebas son las más eficientes, rápidas y confiables para implementar en los procesos de software, dependiendo del estado del código determinan el paso a paso de la prueba y por consiguiente el alto porcentaje de acierto de los resultados.[6]

2.1 Tipos de pruebas automatizadas

2.1.1 Unitarias

Este tipo de pruebas consiste en probar de forma individual los métodos y/o funciones de todos los programas realizados en el proyecto, son consideradas unitarias debido a que son pruebas que verifican el método correcto donde todos sus componentes sean correctos y el resultado sea el esperado, esto es realizado para cada uno de los métodos de forma independiente. [6]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Suelen realizarse en la fase de desarrollo de las aplicaciones de software, ya que al analizar secciones del código permiten la detección de los errores en cada fase de avance evitando que salgan en el producto final, por lo que al realizarlas de manera frecuente suponen un ahorro de tiempo y dinero. También son consideradas para refactorizar el código asegurando garantías de que el módulo programado siga funcionando correctamente.[7]

Las pruebas unitarias deben cumplir con la regla de las 3 A's:

Tabla 1. Regla de 3A's [Propia]

ARRANGE	ACT	ASSERT
Define los requisitos previos que debe tener el código.	Ejecuta el test configurado basándose en los requisitos y listo para el análisis de resultados.	Comprueba los resultados obtenidos si son los esperados o no, decidiendo continuar si es correcto y si no corregir la causa del error.

2.1.2 Funcionales

El tipo de pruebas funcionales corresponde a la verificación de los resultados de los procesos para garantizar que las funcionalidades del software son las correctas y que funcionan sin ningún problema según lo planteado inicialmente.[8]

Principalmente se caracterizan por centrarse en comprobar que los sistemas desarrollados basándose en los requisitos previos funcionen de manera correcta que cumplan con el principal objetivo que se planteó Para así realizar la retroalimentación.[9]

2.1.3 Pruebas de Integración

Verificar los distintos módulos que son utilizados para evaluar el correcto funcionamiento de las aplicaciones o sistemas software con las distintas interacciones que se puedan presentar, por ejemplo, el enlace con base de datos o asegurar que los servicios adjuntos funcionen de manera correcta y según lo esperado, requieren varias partes de la aplicación estén en funcionamiento para dar un diagnóstico específico siendo así las más costosas [10].

2.1.4 Pruebas de punta a punta

Este tipo de pruebas son las correspondientes a evaluar el entorno que tiene el usuario al usar la aplicación, verifican que los flujos que tiene el usuario a disponibilidad trabajen como se espera,[6] por ejemplo:



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Iniciar sesión correctamente
- Cargar enlace web
- Verificación envío email
- Pagos en línea, entre otras.

El propósito principal de estas pruebas es llevar a cabo pruebas que identifiquen las dependencias del sistema y garantizar que el flujo de información se realice de forma correcta.[11]

2.1.5 Pruebas de Regresión

Son pruebas que se realizan el momento de llevar a cabo un cambio en el sistema, para la corrección de errores como para la mejora del código, esto con el fin de asegurar que no se introduzcan errores que comprometan el funcionamiento de otras pruebas ya estabilizadas.[12]

Si al realizar este tipo de pruebas alguna falla es señal de que la nueva actualización ha afectado otra funcionalidad por lo cual es recomendable que hasta que las pruebas de regresión sean totalmente correctas, no se realice ninguna actualización al código fuente.[6]

2.1.6 Pruebas de humo

Son pruebas que aseguran el correcto funcionamiento de las principales características de los programas, o sitio dando aval de la correcta operabilidad, por lo cual es una de las pruebas que es recomendable ejecutar cada día o en cada entorno.[6]

Si la compilación falla en el escenario donde se aplique se debe investigar más a fondo donde se encuentra el error y corregirlo antes de poder realizar otra actualización para evitar errores graves en el resultado esperado.[13]

2.1.7 Pruebas de aceptación

Son pruebas formales necesarias para verificar si el sistema cumple con todo lo que en un principio se dio como requisito, para ejecutarlas es necesario que la fase de desarrollo haya finalizado. Permiten verificar que las características finales estén alineadas con los criterios iniciales y ser aceptadas finalmente.[6]

2.1.8 Pruebas de rendimiento

Comprueban el comportamiento del sistema cuando está sometido a cargas de alto nivel para así medir la fiabilidad, estabilidad y disponibilidad analizando los tiempos de respuesta a ver si son cortos incluso en situaciones complejas.[10]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



3. SELENIUM

Es una herramienta de código abierto que permite la automatización de pruebas a través de navegadores web tales como Mozilla, Chrome, internet Explorer, etc, para así poder interactuar con las páginas y los elementos que tiene cada una de ellas.[14]

Este entorno se utiliza para ayudar a comprobar si las pruebas se desarrollan de manera correcta siendo parte del código que controla el robot, esta herramienta tiene la facilidad de que al escribir el código se pueda autocompletar permitiendo la programación rápida, también puede referenciar a los objetos por medio de Nombres ID, Xpath, tiene funciones globales por defecto que al ser complementadas permite que se desarrolle el código de manera concisa, tales como enviar texto, dar click, enviar secuencias del teclado. Etc. [15]

4. BASES DE DATOS SQL Y ORACLE

4.1 Oracle

Es una herramienta de la relación cliente/servidor que permite la gestión de bases de datos con la funcionalidad de controlar y administrar grandes volúmenes de datos para evitar la pérdida de información y la mala organización. Para ello, hace uso de una distribución gráfica y cómoda de usar, alta seguridad, consultas sencillas y fáciles de entender, usa modelo relacional para visualizar los datos en tablas, etc. [17]

4.2 SQL

Es una de las bases de datos más usadas ya que usa un modelo relacional con lenguaje de consulta estructurado para la gestión de datos en gran escala donde se pueden asignar permisos en tablas procedimientos y vistas, también ejecutar consultas, recuperación de datos, crear nuevas tablas, bases de datos y a su vez insertarle datos, actualizarlos o eliminarlos.[18]

A continuación, se muestra las acciones más importantes y usadas que se pueden realizar a través de la base de datos de SQL, entre las que se destacan las consultas y el manejo de los datos, insertar, actualizar o borrar, que a nivel empresarial es de lo más importante, el manejo de los registros.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Ilustración 2. Que permite hacer SQL [18]

5. LENGUAJE C#

También conocido como “C Sharp” es uno de los lenguajes de programación moderno que basado en lenguajes C, C++, Java y JavaScript, el cual es orientado a objetos proporcionando construcciones naturales para crear y usar componentes de software, permitiendo así la creación de variedad de proyectos como aplicaciones seguras y solidas. [20]

Este lenguaje adaptó también la creación de código en toda la tecnología de la plataforma NET tanto en Framework como Core haciéndolo más polivalente y fácil de aprender sin perder el potencial que adquirió originalmente. [21]

- Puede acceder a el bajo nivel de sistemas operativos
- Trabaja con punteros a memoria
- Interactúa con elementos físicos de los dispositivos como puertos.
- Crea aplicaciones de forma sencilla
- Elimina los elementos que no son necesarios dejando únicamente los importantes.
- Incorpora un mecanismo para el acceso correcto a los datos evitando errores difíciles de arreglar
- Usa bases de datos para interactuar en futuras codificaciones.
- Fácil manejo debido a la compatibilidad que tiene con otros lenguajes.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



La plataforma de desarrollo donde se puede utilizar ese lenguaje, es conocida como Visual Studio totalmente gratuita bajo el entorno .NET en todas las versiones, ya que ha sido diseñado exclusivamente para ello y soporta toda la carga que deriva el uso del mismo. [21]

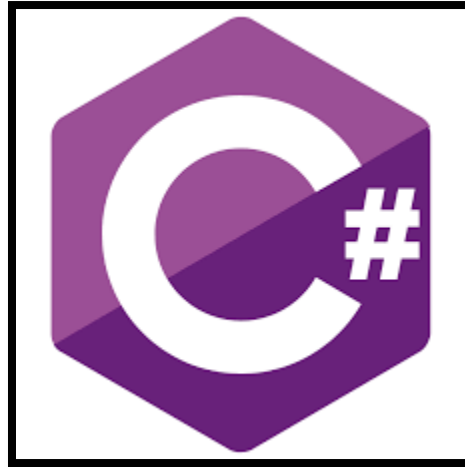


Ilustración 3. C# [22]

6. NET FRAMEWORK

Es un entorno de ejecución que administra aplicaciones que incorporan dos componentes como: Common Language Runtime (CLR) y la biblioteca de clases, donde almacena un código ya probado y reutilizable donde se pueden llamar para la creación de propias aplicaciones.

- Proporciona servicios que permitan administrar la vida útil de los objetos como la liberación de memoria a través de CLR.
- NET Framework define los tipos básicos de compilación en su mismo sistema, y son comunes en todos los lenguajes.
- Contiene cantidad de librerías accesibles en todo momento para simplificar la tarea del programador en la creación de código.
- Incluye bibliotecas específicas para el desarrollo de aplicaciones web ASP.NET y para acceso de datos ADO.NET, aplicaciones orientadas a servicios y aplicaciones de escritorio Windows.
- Interoperabilidad de lenguajes
- Compatibilidad entre todas las versiones



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Permite la ejecución de varias CLR en un mismo equipo sin generar inconvenientes entre ellos. [23]

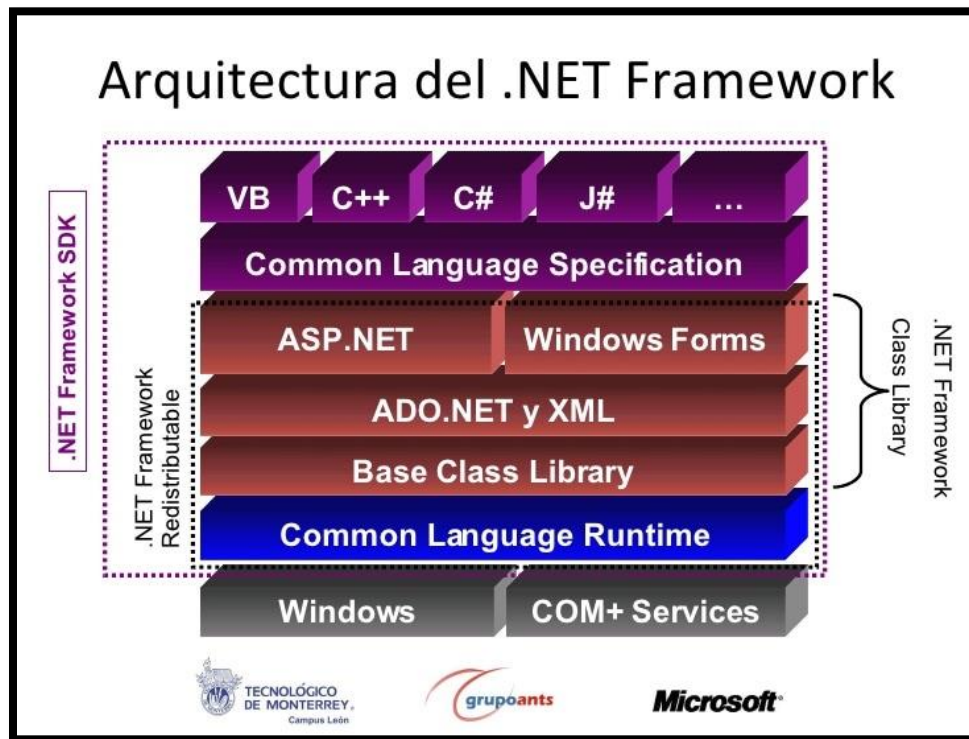


Ilustración 4. Arquitectura NET Framework [24]

La arquitectura de NET Framework se basa en las capas mostradas en la imagen anterior, donde como base se observan el sistema operativo donde también tiene presencia los servicios, actualmente la total disponibilidad de NET esta es en Windows per hay un proyecto llamado MONO para llevarlo a otros sistemas operativos.

En la siguiente capa, se observa el CLR donde se presentan los servicios fundamentales para la correcta ejecución de NET como lo es administrar la memoria, recolección de objetos no utilizables, multiprocesos entre otros, controlando todas las aplicaciones.

Siguiendo con el modelo, se presenta la capa de bibliotecas de clases comunes, donde se almacenan gran número de clases agrupadas en más de 3 ensamblados que permiten el despliegue de ventanas, lectura de archivos, envió de HTML, define los tipos de datos como el objeto base, entre otros.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Para la cuarta capa se presentan las librerías especializadas en la conexión con las bases de datos conocida como ADO.NET y su interacción con ellas gracias a XML para realizar todo de manera transparente.

En la quinta capa se dan las herramientas que permiten la creación de aplicaciones de internet e intranet en la parte ASP. NET y Windows Forms donde se crean aplicaciones para el escritorio de Windows.

Para la capa seis se definen las reglas básicas y comunes que deben cumplir todo lo codificado como tipos de datos, propiedades, métodos, como se deben heredar y la posible estructura de los miembros, intercambiando la información entre lenguajes de programación de .NET.

Finalmente se da la capa de lenguajes de programación donde en Visual Studio se pueden contener más de 30 lenguajes de programación tales como C#, F#, J#, C++, Delphi, Visual Basic, entre otros. [25]

7. NET CORE

Es otra plataforma de desarrollo de Microsoft de código abierto, funcional en cualquier plataforma y de alto rendimiento den la creación de aplicaciones de todo tipo, ya que no tiene dependencia total del sistema Windows, sino que usa un sistema conocido como NuGets donde se pueden subir actualizaciones más seguidas.[26]

- Es multiplataforma, funciona en cualquier sistema operativo
- Permite el funcionamiento en alto rendimiento
- Es open source
- Permite realizar aplicaciones web
- Tiene compatibilidad con contenedores Docker, Azure, entre otras
- Se usa actualmente para el desarrollo de inteligencia artificial, internet de las cosas y desarrollo de juegos.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

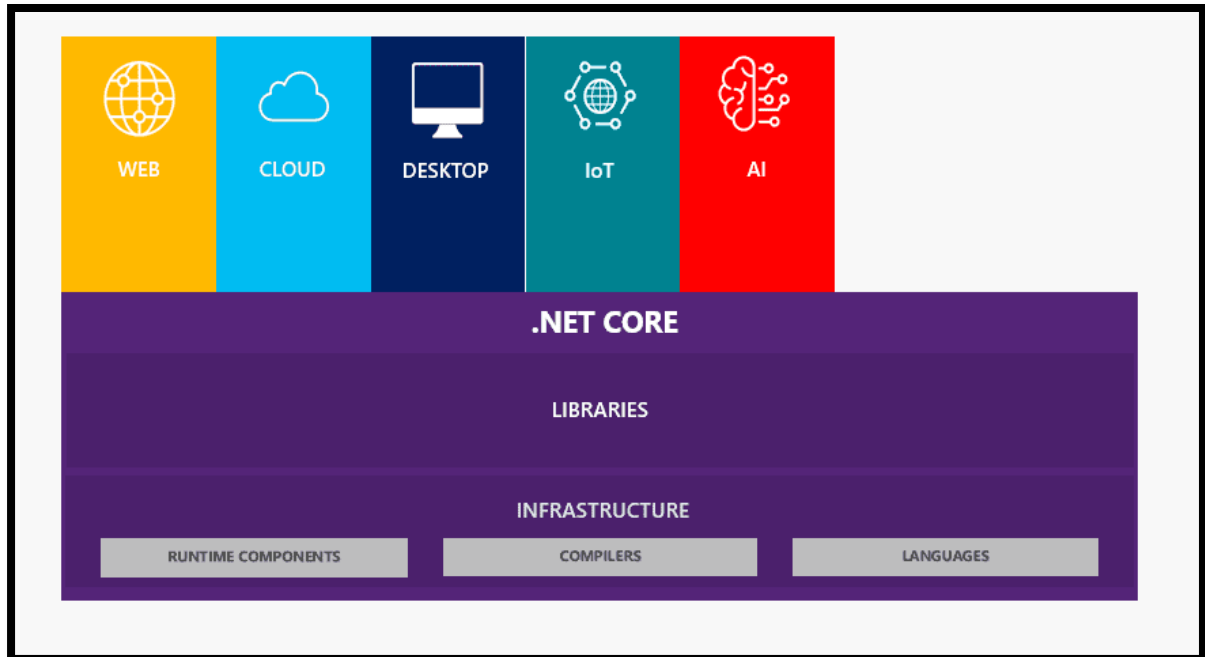


Ilustración 5. Estructura NET CORE [27]

En la imagen anterior se observan los componentes que forman la estructura de NET CORE actual, donde se destacan en la base los componentes Runtime, los compiladores y los lenguajes. Mas arriba se destacan todas las librerías necesarias para el funcionamiento de Core en sus implementaciones, y finalmente se muestra la suite de trabajo donde se destacan las aplicaciones que se pueden realizar a través de NET CORE donde se destacan la inteligencia artificial base primordial del proyecto a desarrollar en este escrito. [27]

8. INTELIGENCIA ARTIFICIAL

La inteligencia artificial es una base donde se realiza la imitación de los procesos de inteligencia humana donde se crean y aplican algoritmos codificados en un entorno dinámico de computación, todo con el fin de entrenar a las máquinas de tal manera que lleguen a pensar y actuar como los humanos.

La IA supone la base de todo el aprendizaje automático debido a que actualmente se manejan cantidades gigantes de información que para el ser humano pueden ser imposible de absorber, interpretar y tomar decisiones basándose en dichos datos, es por eso que ante esta problemática, los ordenadores son mucho más eficientes y programados para ser capaces de manejar



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



información de en grandes cantidades, realizar combinaciones complejas y con un entrenamiento adecuando predecir decisiones como las tomaría un ser humano. [28]

La inteligencia artificial debe cumplir con cinco criterios para ser considerado de utilidad y totalmente funcional:

- Auto gestionables.
- Resilientes.
- Rendimiento mejorable.
- Medibles.
- Aprendizaje continuo y evolutivo. [29]

8.1 Técnicas de inteligencia artificial

- 8.1.1 Machine Learning:** Es la rama que busca el desarrollo de técnicas que permitan a los ordenadores realizar un auto aprendizaje, es decir, aprender por ellos mismos donde se crean programas que permitan generalizar las respuestas a partir de una data previa, y así realizar la predicción de la decisión tomada por el ordenador.
- 8.1.2 Lógica difusa:** Usa un escenario relativo observado como posición diferencial tomando así dos valores al azar para ser contextualizados y relacionados entre sí.
- 8.1.3 Vida artificial:** Estudia las cualidades de los seres humanos a través de una simulación tomando como referencia su entorno de vida.
- 8.1.4 Sistemas expertos:** Permite la creación de sistemas de información con aplicaciones realmente complejas donde se puede configurar la máquina para que de respuestas sobre problemáticas muy específicas.
- 8.1.5 Minerías de datos:** Extrae la información que se encuentra implícita en datos, totalmente desconocida, para poder ser usada en otros proyectos, obteniendo así toda la información que se oculte en ellos.
- 8.1.6 Redes de creencia:** Crea probabilidades al relacionar las variables aleatorias en un modelo pirobalística haciendo uso de grafos dirigidos para indicar una influencia.
- 8.1.7 Ingeniería de conocimiento:** Genera nuevos conocimientos inexistentes a partir de información de bases de datos haciendo cruce de información ya sean documentos, archivos, etc.
- 8.1.8 Redes neuronales artificiales:** Sistema de interconexión de neuronas en una red relacionadas para crear una respuesta de salida ante la presencia de variables de entrada.
- 8.1.9 Sistemas reactivos:** Sistemas que ante la presencia de errores puede afectar en alto nivel a los usuarios que dependen de su uso, ya que son ejecutadas en tiempo real.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- 8.1.10 Redes semánticas:** Representaciones mediante grafos del conocimiento lingüístico usadas para mapas conceptuales, mapas mentales y demás funciones.
- 8.1.11 Lingüística computacional:** Campo de las lingüísticas que es aplicada en la informática, haciendo posible el estudio y tratamiento del lenguaje modelando la lógica del lenguaje natural desde la programación.
- 8.1.12 Procesamiento del lenguaje natural:** Rama de la ingeniería con el objetivo de la lingüística computacional donde se realizan sistemas para la comunicación entre humano y maquina haciendo uso de lenguajes naturales. [30]

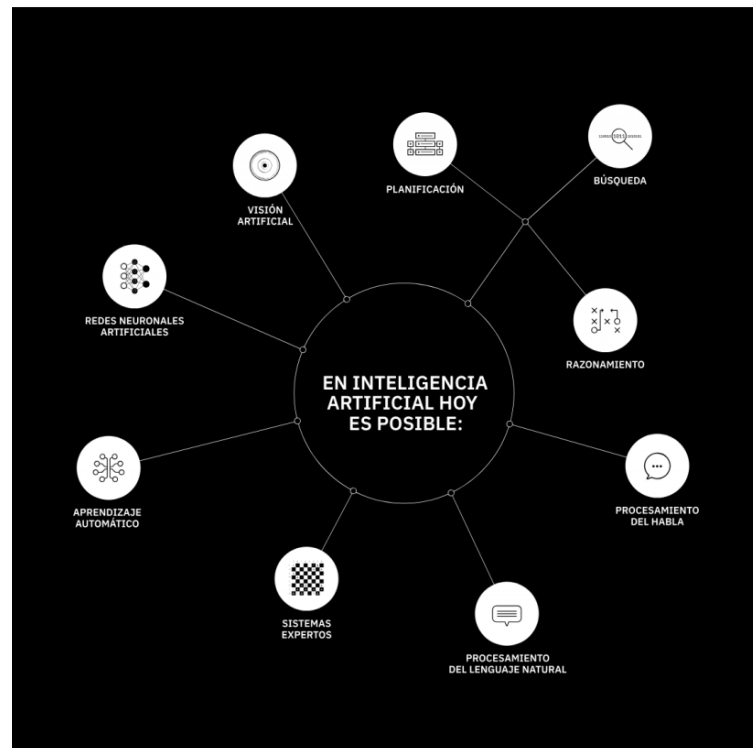


Ilustración 6. Inteligencia Artificial [31]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



9. MACHINE LEARNING

Como se observó en la sección de inteligencia artificial, entre una de sus técnicas sobresale Machine Learning la cual es una de las más importantes debido a su gran alcance y facilidad de crear algoritmos que permitan a las máquinas entender y captar información sin necesidad de una programación previa, por lo tanto, esa va a ser la base de nuestro proyecto y se dará una definición para entenderla un poco más.

Machine Learning es una disciplina del ámbito inteligencia artificial que permite la creación de sistemas predictivos haciendo uso de diferentes algoritmos y técnicas, en la cual la máquina pueden aprender automáticamente al identificar patrones en datos prediciendo los comportamientos futuros, facilitando la operativa humana ya que entrega resultados sin necesidad de intervención humana. [32]

El funcionamiento de Machine Learning se basa principalmente en los algoritmos ya que a través de él se define el contexto a usar y que es lo que debe aprender o no la máquina, de donde saca la información para su entrenamiento, y mucho más. Al tener este algoritmo programado con los parámetros necesarios, la máquina procede a realizar sus propios cálculos programándose a sí mismas para así diseñar sus predicciones, guardarlas y al momento de evaluar el sistema creado tenga la capacidad de dar una respuesta correcta. [33]

El comportamiento de esta tecnología se basa en cinco etapas primordiales que siempre se cumplen a cabalidad para tener un aprendizaje automático de correcto funcionamiento, como se muestra a continuación:

1. Obtiene la información con la que se entrenará el algoritmo.
2. Limpia, prepara y manipula la información.
3. Entrena el modelo basándose en la información previa.
4. Se hace el testeo de cómo se comporta el modelo creado.
5. Finalmente se evalúa y si es necesario, mejorar el algoritmo.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

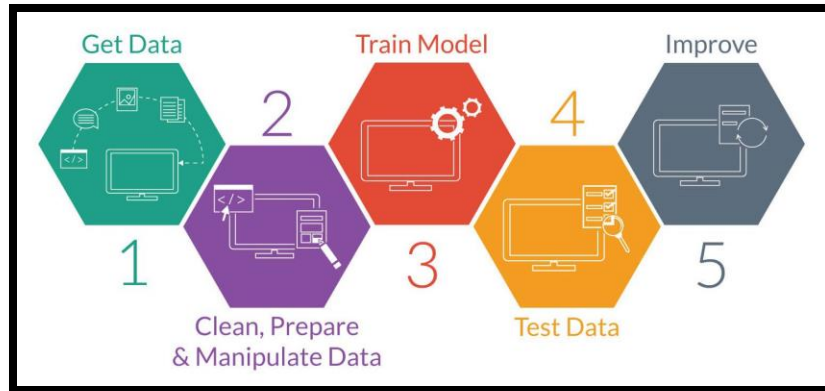


Ilustración 7. Comportamiento Machine Learning [34]

9.1 Tipos de Machine Learning

Machine Learning al ser una técnica de IA tan amplia, presenta varios tipos, como se muestra a en la siguiente imagen, los cuales se explicarán de manera detallada a continuación.

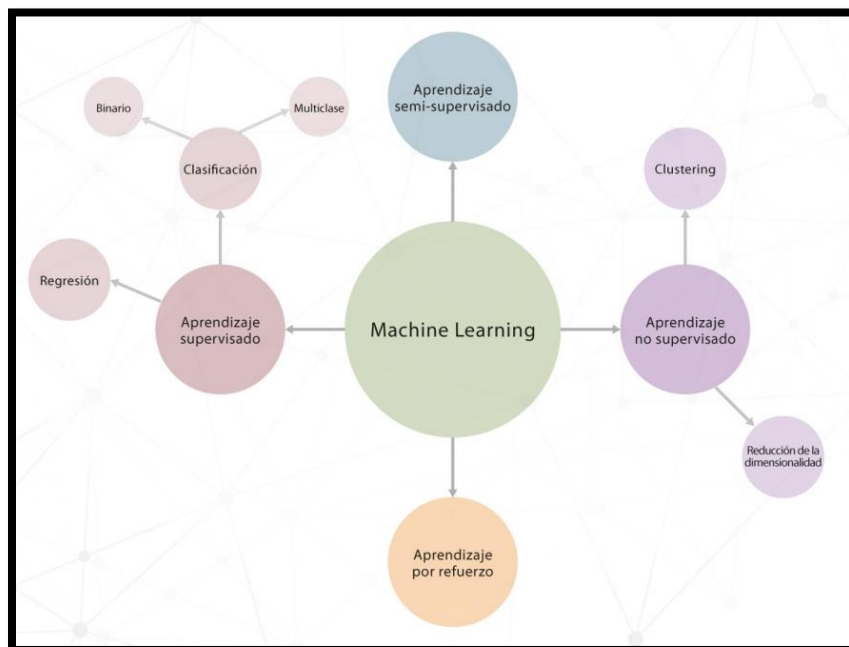


Ilustración 8. Tipos de Machine Learning [34]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



9.1.1 Aprendizaje Supervisado

Este tipo de aprendizaje se caracteriza principalmente por tener en cuenta datos etiquetados, es decir, etiquetas para conocer qué es lo que debe aprender, siendo uno de los aprendizajes más usados en ámbitos de clasificación. Para el desarrollo de este aprendizaje se dan dos modelos de suma importancia como lo son: [34]

- **Clasificación:** Tiene como objetivo predecir las etiquetas de clase por categoría de registros nuevos en base a lo aprendido con el entrenamiento al que fue creado el modelo.

Esta etiqueta dependiendo del caso puede ser binaria o multiclase dependiendo de lo que se quiere predecir; para ser binaria si solo existen dos estados de la etiqueta mientras que si existen más de dos clases es conocida como multiclase [35].

Para entender un poco más esta clasificación binaria y su comportamiento en cuanto a matemática y probabilidad, se tiene la siguiente imagen donde se tienen las variables X_1 Y X_2 las cuales representan las características de las dos clases presentes, como se muestra en la ilustración 11.

Teniendo estas dos variables, esta tecnología realiza la relación entre ellas para encontrar un punto en el cual se pueda trazar la línea de separación generando así una puntuación de predicción. [59]

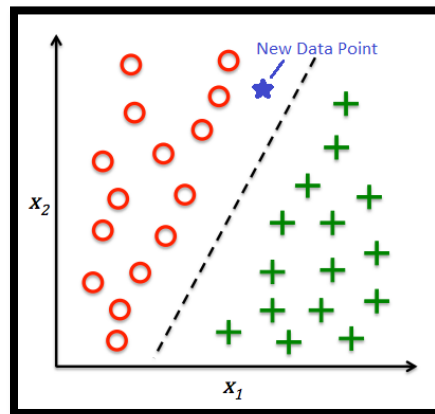


Ilustración 9. Clasificación [36]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



La puntuación mencionada anteriormente, se comporta como la ilustración 12, allí se compara con el umbral que determina en que sección se encuentra la clasificación positiva (parte derecha de la gráfica) o negativa (parte izquierda de la gráfica). Al ser predicciones de puntuaciones se pueden presentar otros dos estados distintos al positivo o negativo, conocidos como falsos positivos y falsos negativos, estos tienen una zona especial correspondiente a la zona con líneas rojas, evitar la presencia de las puntuaciones en esta zona se basa en la excelente elección del algoritmo y se puede conocer haciendo uso de métricas, las cuales se explicarán más adelante. Las puntuaciones de verdaderos positivos y negativos se encuentran en los picos altos de cada separación y más alejadas del límite del umbral, apoyada en el nivel de precisión que del modelo. [58]

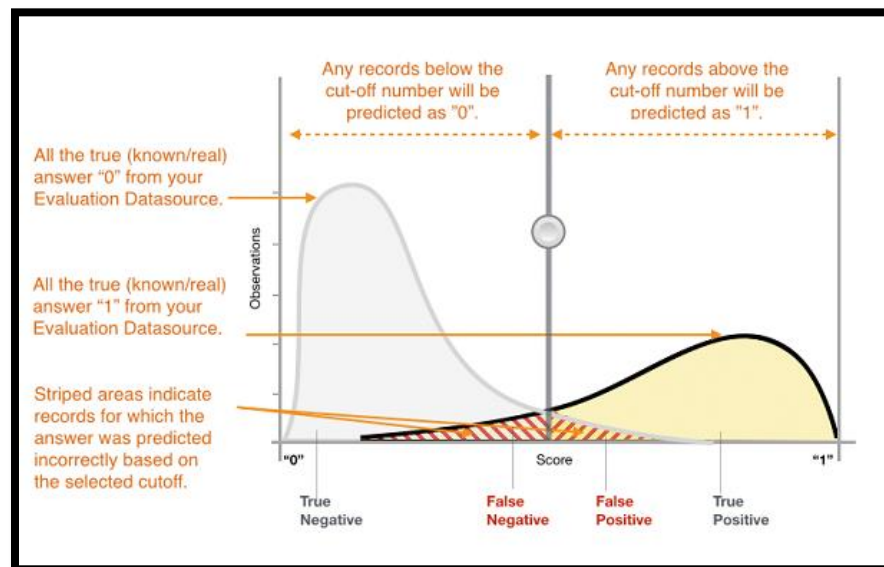


Ilustración 10. Comportamiento matemático de clasificación binaria [58]

El análisis anterior se basa en la evaluación del algoritmo basándose en elementos muy importantes en el machine Learning como lo son las **métricas**, las cuales permitirán realizar un barrido de lo que se presentó en la clasificación binaria creada, por ello se hace uso de las siguientes:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Matriz de confusión

- Indica la cantidad de verdaderos positivos o negativos y los falsos positivos o negativos.

		predicción	
		0	1
realidad	0	TN	FP
	1	FN	TP

Precisión

- Mide la calidad del modelo al ejecutar la predicción con la razón entre verdaderos positivos y la suma entre ellos más falsos positivos.

$$precision = \frac{TP}{TP + FP}$$

Recall

- Cantidad de predicciones que es posible identificar el modelo.

$$recall = \frac{TP}{TP + FN}$$

F1

- Permite comparar el rendimiento del modelo haciendo uso de Recall y precisión.

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Accuracy

- Porcentaje de acierto de las predicciones.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- TP (Positivo verdadero) FP (Positivo falso) TN (Negativo verdadero) FN (Falso negativo). [60]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Una de las aplicaciones más comunes donde se usa este tipo de Machine Learning, específicamente clasificación binaria, es en la detección de correos spam, ya que solo maneja 2 clases como lo son el sí o el no, entonces el algoritmo recibe el correo electrónico y basándose en las características aprendidas predice si pertenece a spam o no-spam.

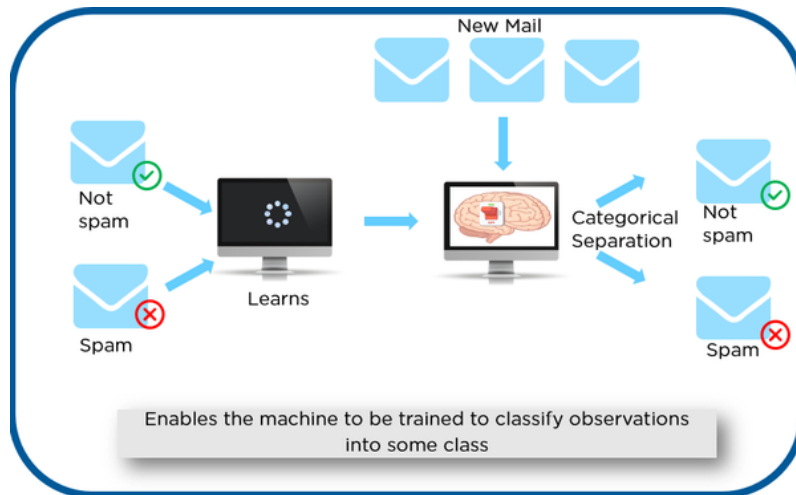


Ilustración 11. Clasificación Binaria [37]

Para el uso de clasificación multiclase, se comporta de igual manera que la binaria en cuestión de análisis de las predicciones, pero al ser muchas clases presentes, se deben tener en cuenta si es aplicable la técnica de líneas o es necesario hacer uso de otras técnicas de machine Learning, se tiene en cuenta la siguiente imagen, donde se observa el comportamiento de las líneas de separación ante la presencia de varias clases para su separación, si las métricas dan un porcentaje bajo, es necesario hacer uso de nuevas técnicas para mayor precisión.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

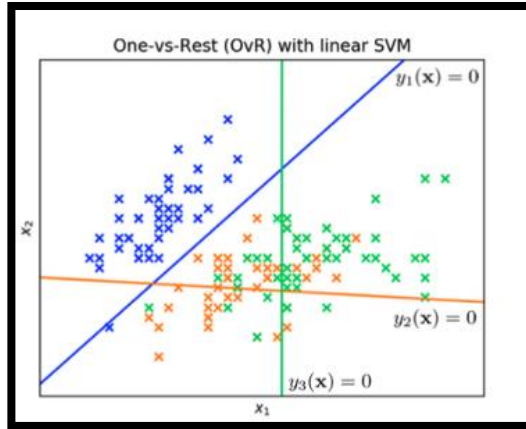


Ilustración 12. Comportamiento clasificación multiclase [59]

Un ejemplo, sería predecir los tipos de perros que existen, en este caso se hace uso de 4 clases que corresponden a las etiquetas Pitbull, Beagle, Pastor Alemán y Chihuahua, al entrenarse el algoritmo e ingresarle la imagen permita clasificarlos según las razas.

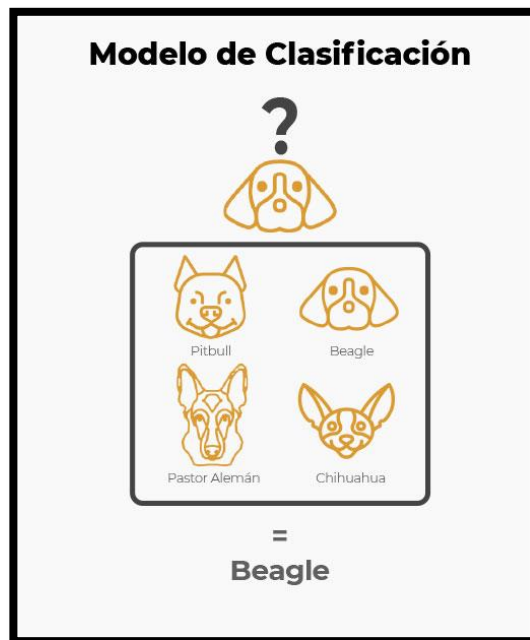


Ilustración 13. Clasificación multiclase [38]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Para realizar clasificación de 2 o más clases, se destacan varias técnicas que permiten la realización de este tipo de machine Learning, como las siguientes:

- **Regresión Logística:** Es una técnica que simula una red neuronal con la diferencia que solo hace uso de una neurona para realizar su predicción, es un método muy simple, efectiva, fácil de interpretar para muchos problemas que necesitan clasificación. [39]

Esta técnica se basa en la función de proporción de probabilidades, donde la letra P corresponde a probabilidad de una predicción verdadera, $Probabilidades = \frac{P}{1-P}$, que permite conformar la formula principal para expresar las relaciones lineales entre los variables de características y el algoritmo de funciones impares,

$$\text{Log}(P) = \log \frac{P}{1-P}$$

$$\text{Log}(P(y = 1|x)) = W_0X_0 + \dots + W_mX_m$$

Donde en la probabilidad se ingresa a que clase, este caso 1, y las características X, pero para realizar la probabilidad de que la predicción realice la clasificación se aplica la inversa a la ecuación de Log, dando como resultado la **función sigmoide**, como la siguiente formula, y se comporta como la ilustración 14, dando un valor entre 1 y 0 siendo probabilidades.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

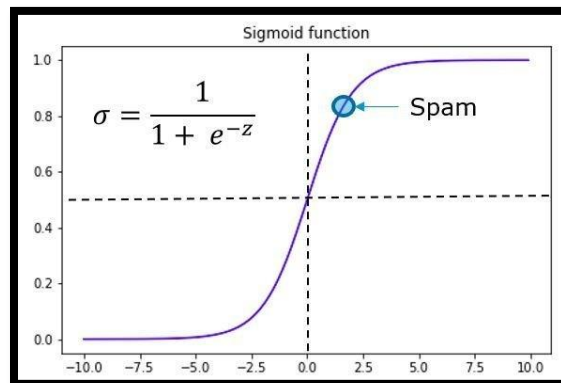


Ilustración 14. Función Logística [39]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Para entender ver gráficamente el proceso, se hace uso de la siguiente imagen, en ella se muestran unas series de variables que corresponden a lo siguiente:

- Los valores de X corresponden a los distintos atributos que presenta el problema, en el ejemplo del correo spam, los valores de X serían cuantas veces aparece cada palabra en un texto y así realizar la predicción.
- La predicción resultante corresponde a la variable Y.
- La primera función corresponde a una combinación lineal, sacada de aplicar la función de logit anteriormente.
- La segunda función corresponde a una función logística, en la cual se representa de la forma mostrada en la ilustración 14, en ella se observa el comportamiento que tiene donde se destacan: Acotada entre 0 y 1, Sus resultados se interpretan como probabilidades, 85% sería 0.85 y e clasificación binaria se clasifican con probabilidades mayores a 0.5 o menores a 0.5
- Las variables W hacen referencia a los coeficientes con los que se combinan los datos de entrada formando la línea. [39]

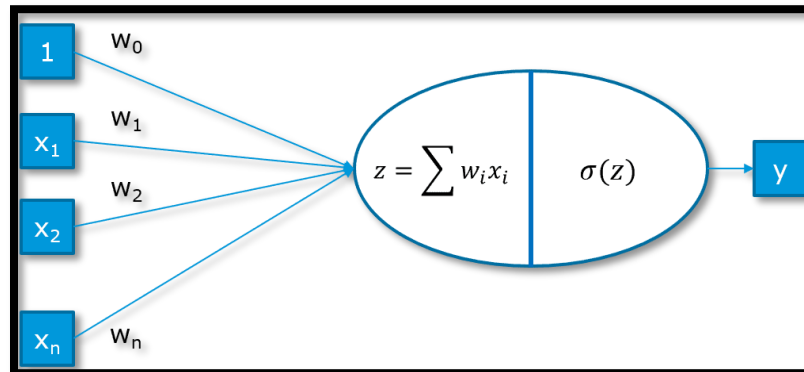


Ilustración 15. Modelo regresión lineal [39]

- **Vectores de Soporte (SVM):** Es una técnica de machine Learning para clasificación que hace uso de una línea que maximiza el margen de separación entre las clases presentes, si se habla de clasificación binaria o multiclase, sirve de igual manera. [40]
Para realizar el margen máximo se establece lo siguiente, a nivel matemático:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

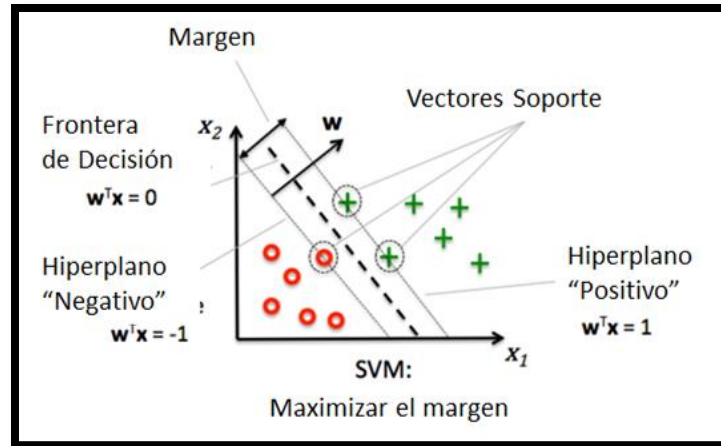


Ilustración 16. Maximizar margen [60]

Primero se deben establecer dos líneas paralelas a la línea de decisión original, en la ilustración anterior corresponden a las líneas de color gris, entre más ancha sea la distancia entre ellas, menos es el error de predicción que se puede presentar, la distancia del margen se detiene hasta encontrar los vectores de soporte de diferentes clases, permitiendo que se detecten las zonas de las clases presentes, en este caso sería el positivo y negativo, conocidos como hiperplanos.

De manera matemática, estos hiperplanos se comportan de la siguiente manera, donde se basa en los coeficientes con cada atributo necesario para predecir, en sumatoria.

$$W_0 + W^T X_{pos} = 1$$

$$W_0 + W^T X_{neg} = -1$$

Y al restarse entre ellas y haciendo uso del vector normalizado conocido como W , se obtiene la función que da como resultado el margen que se requiere:

$$\frac{W^T (X_{pos} - X_{neg})}{\|W\|} = \frac{2}{\|W\|}$$

La parte izquierda corresponde al tamaño del margen maximizado. [60]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Como se observa, en la siguiente ilustración, otra forma de aplicar vectores y hallar el hiperplano con mejor clasificación de las clases, se puede usar el truco del Kernel, este consiste en crear una dimensión nueva que permita la separación de clases, es decir pasar de un plano XY a uno XYZ, que no sean posibles de manera lineal, para ello se trazan líneas horizontales que separen los datos ubicados en el valle y pico de la gráfica normalizada, como se muestra en la ilustración 18, separando así los hiperplanos. [40]

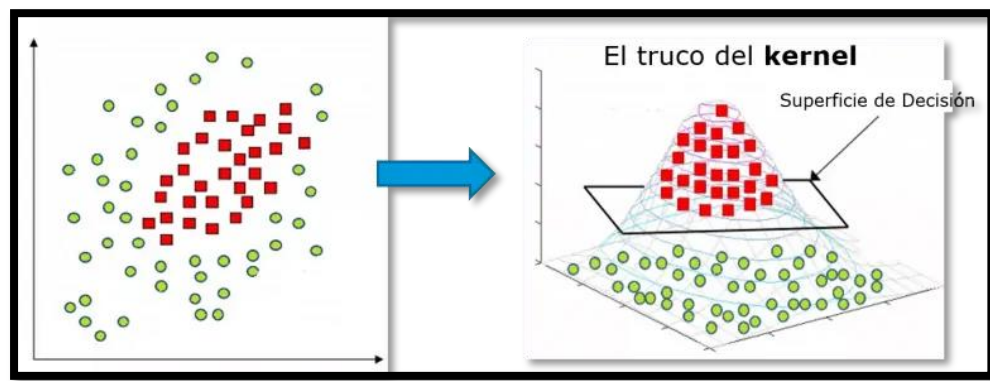


Ilustración 17. Truco Kernel [40]

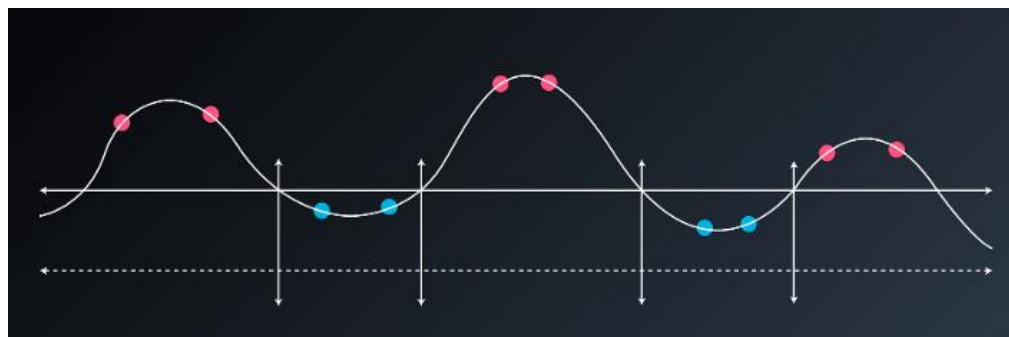


Ilustración 18. Separación Hiperplanos [60]

- **Árbol de decisión:** Técnica de Machine Learning para clasificación que permite la separación de datos teniendo varias clases con características específicas, haciendo uso de preguntas en secuencia hasta encontrar la predicción adecuada, siguiendo el modelo de la siguiente imagen:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co

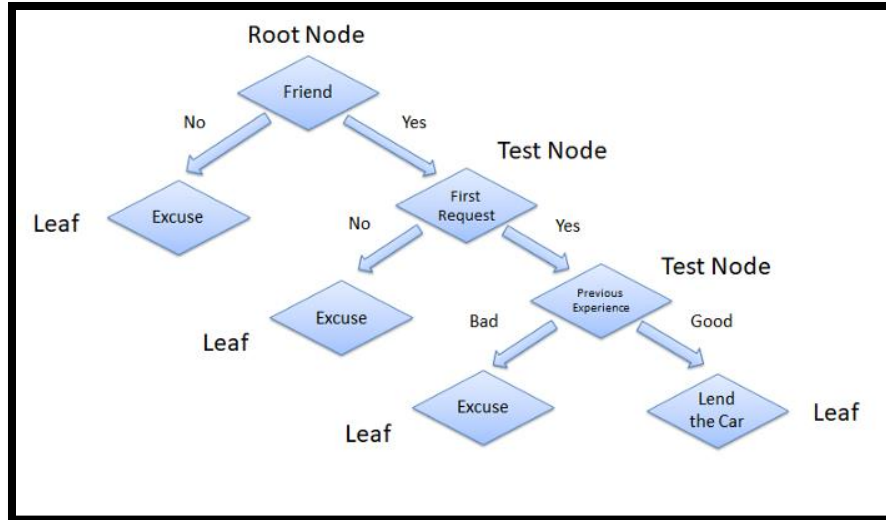


Ilustración 19. Árbol de decisión [60]

Como ejemplo, la ilustración 19, crea un árbol de decisiones basándose en si es recomendable o no prestarle el auto a un amigo, por lo cual el algoritmo toma en cuenta 3 ramas y 6 hojas, ante una respuesta de porque no o porque si, creando así nodos que al final le permitan tomar la mejor decisión, que en este ejemplo es un si debido a las previas experiencias con el [60]

- **Regresión:** Es otro modelo de Aprendizaje supervisado en el cual se predice un resultado numérico dentro de infinidad de posibilidades, es decir ayuda al desarrollo de algoritmos que permitan clasificar por ejemplo cuanto va a vender una inmobiliaria basándose en infinidad de datos. Por ejemplo, cálculo de probabilidades de tener un buen beneficio con respecto a la inversión que se realiza. [42]

Su comportamiento se puede evidenciar de las siguientes maneras:

- **Regresión Lineal:** Es una técnica paramétrica donde se saben con anticipación cuantos parámetros se van a usar, estos equivalen a los coeficientes, todo esto con el fin de encontrar los mejores parámetros para los datos, entre menor sea las posibilidades de error más asertivo será la predicción. [43]



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

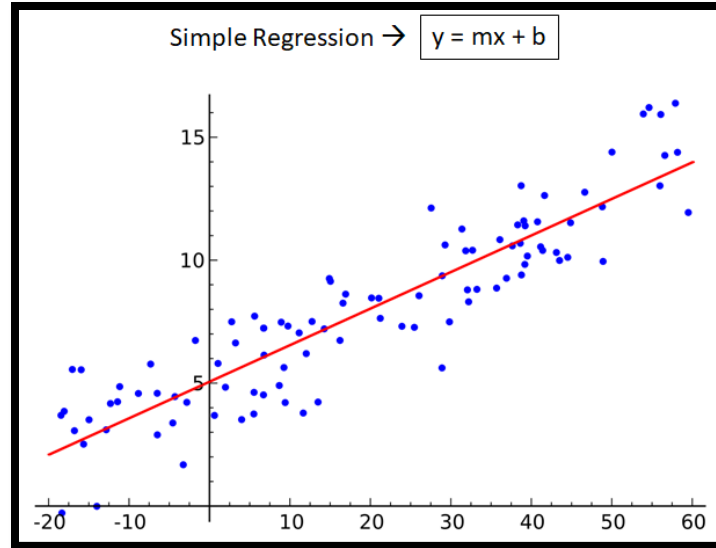


Ilustración 20. Comportamiento de Regresión Lineal [61]

De manera matemática, la regresión lineal se basa en probabilidades que indiquen la tendencia que toman un conjunto de datos continuos conocidos como características ante un resultado final, conocido como variable independiente o predictora.

La fórmula en la que se basa el comportamiento de Regresión Lineal es la siguiente:

$$Y = wx + b$$

Donde el correcto proceso de aprendizaje se deriva en estimar los parámetros w y b con los mejores valores posibles, para que la línea creada tenga el porcentaje de error mínimo y el mayor acierto. [61]

Para la medición de regresiones lineales, se hace uso de métricas en las que se destacan:



SC-CER96940

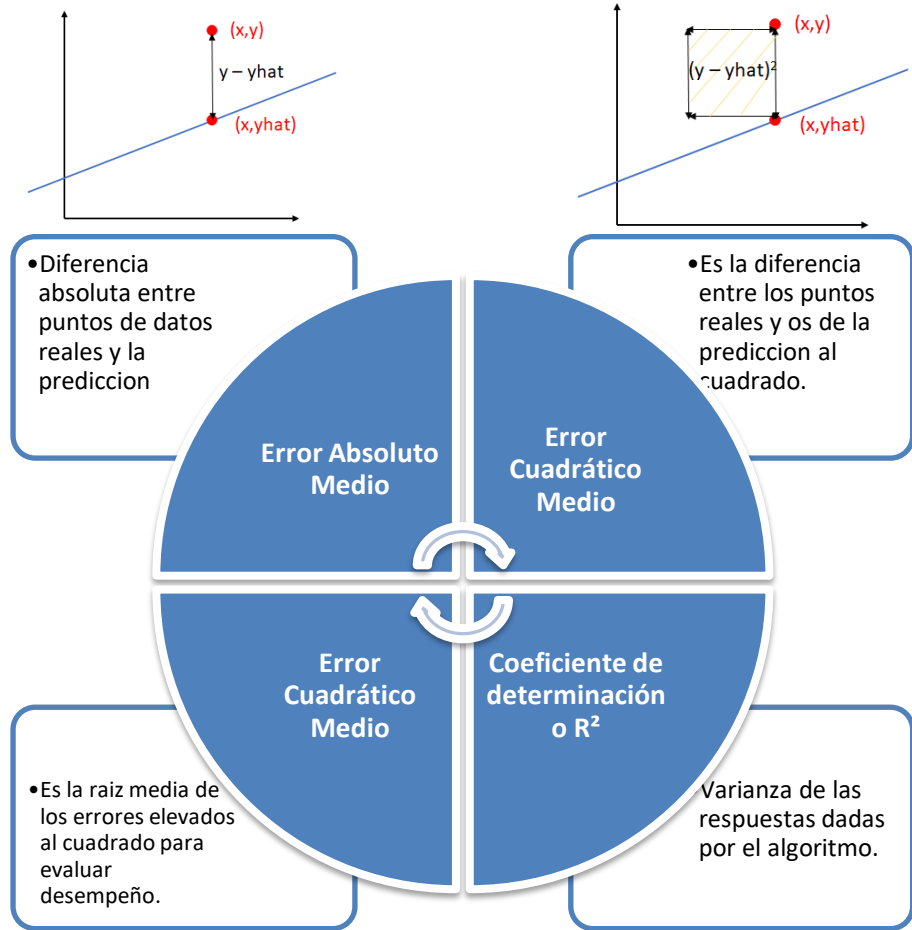


“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$\text{Coeficiente de Determinación} \rightarrow R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$\text{Suma de Cuadrados Total} \rightarrow SST = \sum (y - \bar{y})^2$$

$$\text{Regresión de Suma de Cuadrados} \rightarrow SSR = \sum (y' - \bar{y}')^2$$

$$\text{Error de Suma de Cuadrados} \rightarrow SSE = \sum (y - y')^2$$



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- **Regresión no lineal:** Esta técnica que permite hacer predicciones basándose en variables independientes y los atributos conocidas como variables dependientes haciendo uso de funciones no lineales que permitan acoplarse a la posición que tienen los datos y tener altas probabilidades de acierto, como se muestra en la siguiente imagen, se tiene un conjunto de datos que su posición no permite acoplar la regresión lineal, por lo que se recurre a usar la no lineal basándose en la curva que adquiera la mayor cantidad de datos.

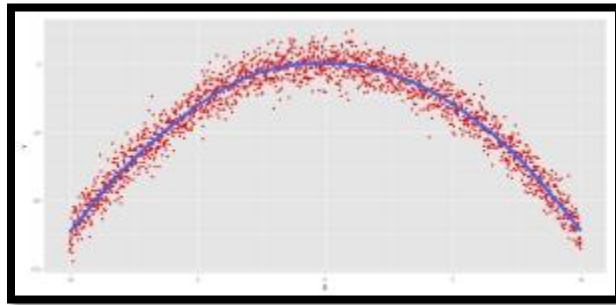


Ilustración 21. Regresión no lineal [61]

Todo con el propósito de obtener los valores de los parámetros asociados que hagan una mejor curva de ajuste. [61]

9.1.2 Aprendizaje no supervisado: Es un tipo de aprendizaje que realiza el entrenamiento y predicciones con datos que no poseen una etiqueta, por lo cual usan técnicas que permitan las agrupaciones de datos con afinidad o similitudes realizado comparaciones y distribuirlos en grupos de acuerdo a los atributos parecidos.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co

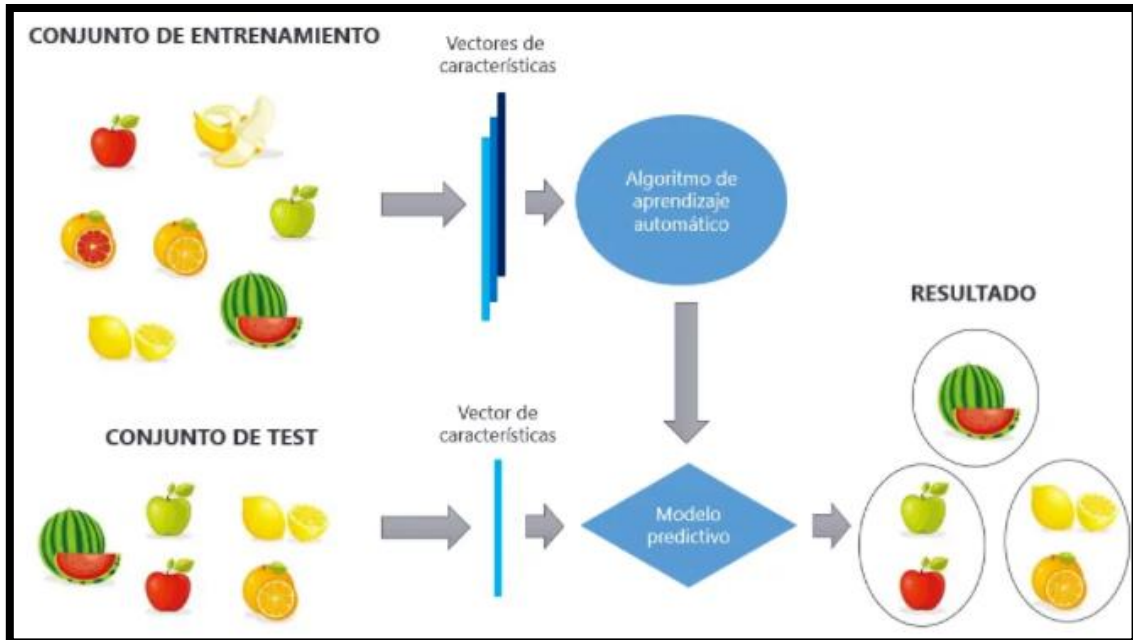


Ilustración 22. Aprendizaje no supervisado [45]

Este tipo de machine Learning usa dos modelos de aprendizaje muy importantes como lo son el Clustering y la Reducción de dimensionalidad, los cuales serán explicados a continuación:

- **Clustering:** Modelo que clasifica los datos haciendo uso de grupos para clasificar teniendo en cuenta las características de las variables, como el color, la forma, tamaño, etc

Por ejemplo, se tiene una canasta con diferentes frutas y se realiza la configuración del modelo de Clustering de tal manera que clasifique esas frutas teniendo en cuenta sus características como lo son su color, tamaño, forma, obteniendo como resultado la siguiente predicción



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

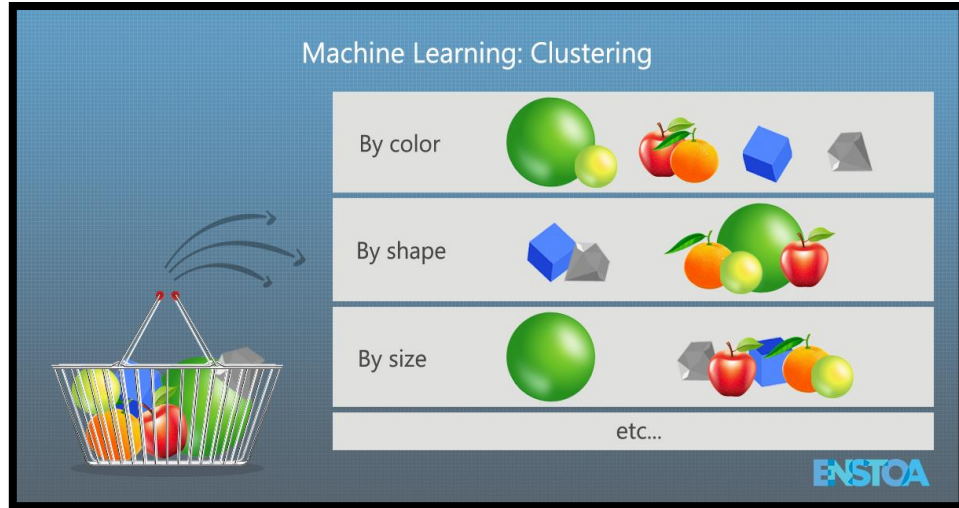


Ilustración 23. Clustering [46]

La mejor técnica se hace pie en este modelo es conocida como cluster k-means, se comporta basándose en la variable K ya que ahí se coloca cuantos grupos queremos usar para la predicción, teniendo esto se basa en centroides de puntos donde los distribuye de la manera que se hagan notar los claramente los grupos.[47]

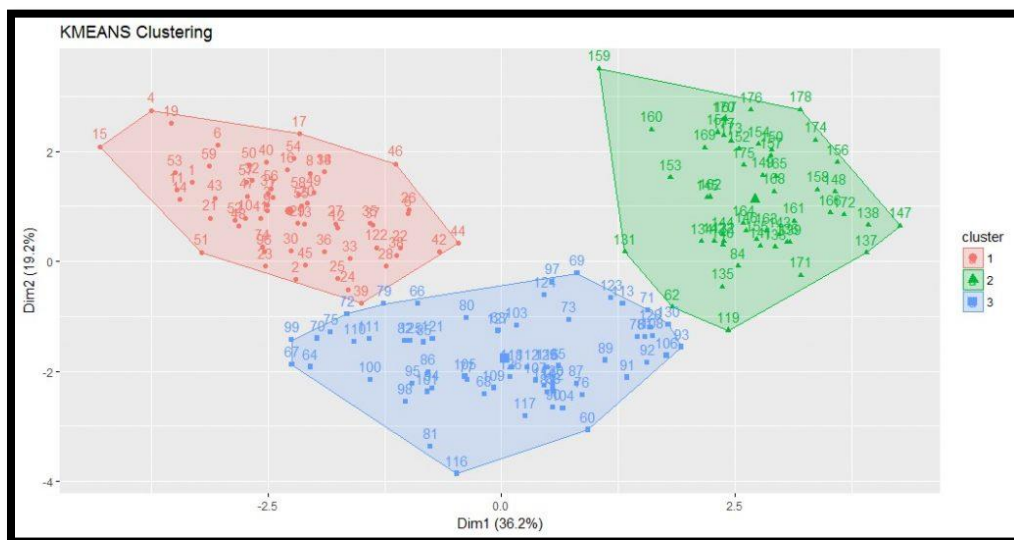


Ilustración 24. Técnica KMEANS [47]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
 Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
 www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- **Reducción de dimensionalidad:** Este modelo tiene como finalidad reducir el número de variables a tener en cuenta para el análisis y en que magnitud de importancia tiene cada una de ellas. [45]

Para entender un poco mejor la definición, se pone de ejemplo, la siguiente imagen:

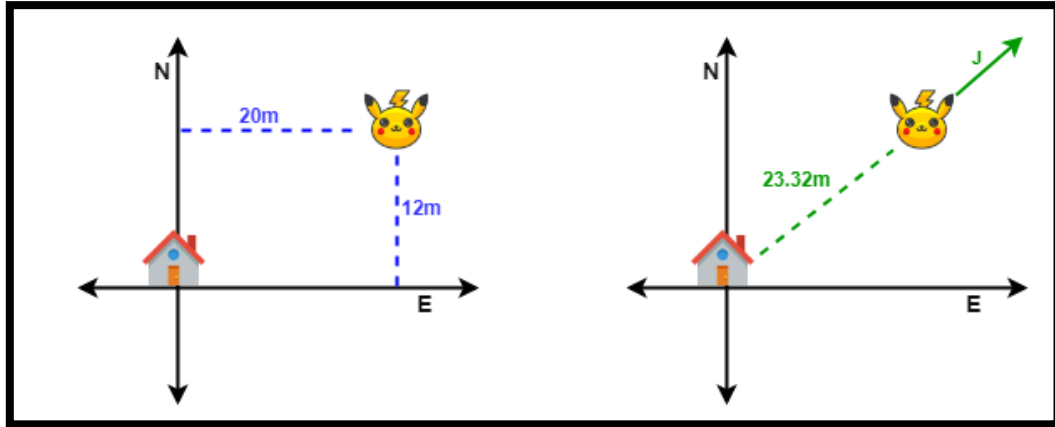


Ilustración 25. Reducción de dimensionalidad [48]

En la parte izquierda, se observa la ubicación de la animación conocida como Pikachu, donde para dar las coordenadas se necesitó de dos variables como lo es distancias en Norte y Este. Mientras que en la parte derecha se encuentra una forma de dar esa misma ubicación necesitando únicamente una sola variable realizando un cálculo matemático entre las dos variables de entrada realizando así la misma función, ese es el objetivo de la reducción de dimensionalidad.

Las técnicas usadas para realizar este modelo se resumen en el siguiente cuadro



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co

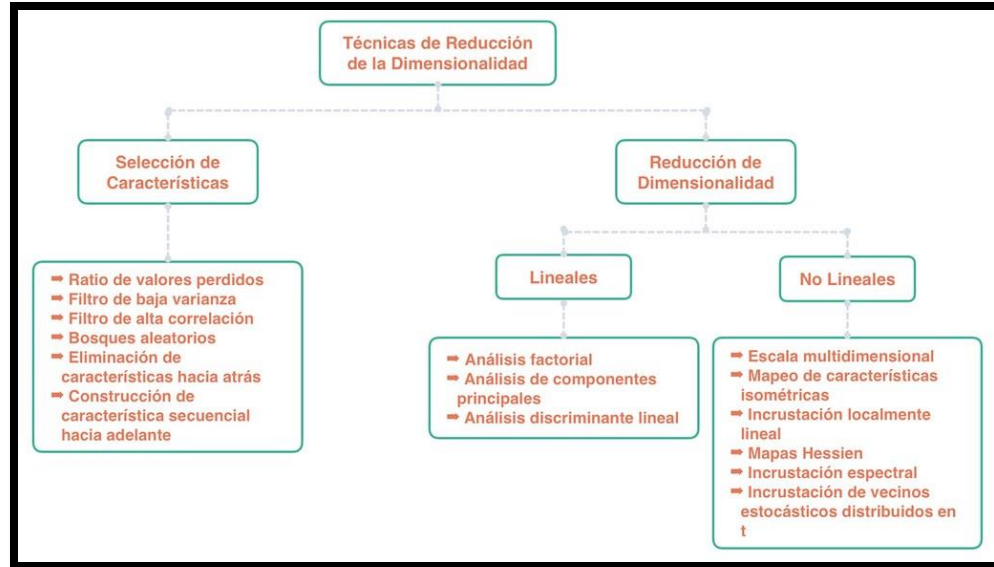


Ilustración 26. Técnicas de Reducción de dimensionalidad [49]

9.1.3 Aprendizaje por refuerzo: El ultimo tipo de Machine Learning es un aprendizaje basado en prueba y error, es decir, el algoritmo aprende basándose en los comportamientos deseados y penalizaciones ante los no deseados logrando así la percepción del entorno en que es lo correcto y lo incorrecto. [34]



Ilustración 27. Aprendizaje por refuerzo [50]



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



Basándose en toda la información presentada acerca de machine Learning y sus modelos, para el desarrollo de este sistema de clasificación se aplicará el aprendizaje supervisado, gracias a la facilidad que da el uso de etiquetas, con el modelo de clasificación, específicamente, la binaria, se hará uso de técnicas más estables como lo es la regresión logística gracias a su fácil entender matemático y alta precisión para este tipo de clasificación.

10. ML.NET (MACHINE LEARNING .NET)

Es conocida por ser una multiplataforma de Machine Learning totalmente gratuita que hace parte de la tecnología de NET. Framework para ser ejecutada en Windows y NET. Core para su uso en demás sistemas operativos. Permite el total uso de la tecnología de inteligencia artificial en el ámbito de aprendizaje automático en las aplicaciones que se crean con .NET permitiendo la programación del algoritmo que recibe los datos para su entrenamiento y evaluación. [51]

Esta plataforma puede usar Machine Learning haciendo uso de sus librerías en la creación del algoritmo a el gusto del programador, unos ejemplos de que se puede hacer con esta plataforma.

Clasificación de comentarios positivos o negativos.

Valores de predicción y regresión para ubicar una vivienda

Detecta anomalías, usado en bancos para detectar fraudes

Recomendaciones para las empresas y sus productos

Clasificación de imágenes, muy usadas en el ámbito de la salud

Y como las aplicaciones anteriores, se pueden realizar muchas más gracias a la facilidad que da ML.NET de aplicarlo a proyectos con toda su compatibilidad con la tecnología NET, ya que esta funciona en la mayoría de los dispositivos actuales, es por ello que los programadores tienen la facilidad de programar a su estilo sin ningún problema. [52]

Esta tecnología es de las mejores en los siguientes modelos de machine Learning vistos anteriormente, como lo son:

- **Agrupación**
- **Regresión**
- **Clasificación**

10.1 Características

- Creado para desarrolladores que manejen la tecnología .NET y todos los lenguajes de programación



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Ofrece herramientas productivas que permitan crear, entrenar e implementar fácilmente los modelos de ML a su gusto de alta calidad.
- Permite el uso de otras bibliotecas de aprendizaje automático como lo es Infer.NET, TensorFlow y ONNX.
- Confiable, ya que se han desarrollado algoritmos de ML para programas como Bing, Outlook, Microsoft Defender. [53]
- Permite entrenar, construir y enviar modelos personalizados en ML usando C# o F#.
- Hace uso de ciertas características y herramientas como ML.NET CLI, ML.NET Model Builder, AutoML facilitando la integración del ML en las aplicaciones.
- Tiene compatibilidad con sitios de Azure en la cual puede realizar modelos en los servicios de nube.
- Puede ser usado en cualquier aplicación de la tecnología .NET son sus aplicaciones, servicios, Docker, etc. [54]

10.2 Componentes de ML.NET

Para iniciar la creación de modelos, en ML. NET se tienen en cuenta 7 pasos principales para considerarlo completo y estable.

Tabla 2. Paso a paso para crear un modelo con ML, NET [54]

Paso	¿Qué se hace?
Crear contexto ML.NET	<p>Es el punto de partida necesario para aplicar todas las operaciones ML.NET, es el ambiente donde se utilizan todos los aspectos de creación del modelo, por lo general, se maneja de la siguiente manera.</p> <pre>var mlContext = new MLContext ();</pre>
Cargar los datos	<p>Estos la entrada de datos que son los utilizados para entrenar el algoritmo, que, al ser aprendido, permitan que el modelo de una predicción.</p> <p>Se representan como un IDataView, donde describe la forma de los datos por ejemplo filas y columnas. Estos datos pueden contener números, texto, valores booleanos, vectores y más.</p> <p>Una forma de cargar datos sería la siguiente:</p> <pre>IDataView trainingData = mlContext.Data. LoadFromTextFile</pre>



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
 www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



	<pre>IDataView trainingData = mlContext.Data. LoadFromEnumerable</pre> <p>El ultimo código es otra forma de cargar datos desde colecciones en memoria como XML/JSON, bases de datos, etc.</p>															
<p>Transformar datos</p>	<p>Es el paso en el que se toman los datos cargados, se trabajan y devuelven datos nuevos y transformados de tal manera que puedan ser usados para entrenar el modelo de ML, para hacer eso, ML.NET hace uso de:</p> <div style="border: 1px solid gray; padding: 10px; margin: 10px 0;"> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 45%; border: 1px solid gray;">Text</th> <th style="width: 10%;"></th> <th style="width: 45%; border: 1px solid gray;">Featurized text</th> </tr> </thead> <tbody> <tr> <td style="border: 1px solid gray;">Wow... Loved this place</td> <td style="text-align: center;">-----▶ FeaturizeText(...) -----▶</td> <td style="border: 1px solid gray;">[0.76, 0.65, 0.44, ...]</td> </tr> <tr> <td style="border: 1px solid gray;">Crust is not good.</td> <td></td> <td style="border: 1px solid gray;">[0.98, 0.43, 0.54, ...]</td> </tr> <tr> <td style="border: 1px solid gray;">Not tasty and the texture was just nasty.</td> <td></td> <td style="border: 1px solid gray;">[0.35, 0.73, 0.46, ...]</td> </tr> <tr> <td style="border: 1px solid gray;">The selection on the menu was great.</td> <td></td> <td style="border: 1px solid gray;">[0.39, 0.00, 0.75, ...]</td> </tr> </tbody> </table> </div>	Text		Featurized text	Wow... Loved this place	-----▶ FeaturizeText(...) -----▶	[0.76, 0.65, 0.44, ...]	Crust is not good.		[0.98, 0.43, 0.54, ...]	Not tasty and the texture was just nasty.		[0.35, 0.73, 0.46, ...]	The selection on the menu was great.		[0.39, 0.00, 0.75, ...]
Text		Featurized text														
Wow... Loved this place	-----▶ FeaturizeText(...) -----▶	[0.76, 0.65, 0.44, ...]														
Crust is not good.		[0.98, 0.43, 0.54, ...]														
Not tasty and the texture was just nasty.		[0.35, 0.73, 0.46, ...]														
The selection on the menu was great.		[0.39, 0.00, 0.75, ...]														
<p>Elegir algoritmo</p>	<p>En este paso, se elige una tarea de ML que se adapte a su escenario donde se pueden escoger entre más de treinta algoritmos para la variedad, entre las que se destacan:</p> <ul style="list-style-type: none"> • Clasificación binaria: <i>AveragedPerceptronTrainer,</i> <i>SdcaLogisticRegressionBinaryTrainer</i> • Clasificación multiclase: <i>LightGbmMulticlassTrainer, OneVersusAllTrainer</i> • Regresión: <i>LbfgsPoissonRegressionTrainer, FastTreeRegressionTrainer</i> • Agrupación o Clustering: <i>KMeansTrainer</i> • Detección anomalías <i>RandomizedPcaTrainer</i> • Recomendaciones: <i>MatrixFactorizationTrainer</i> • Clasificación: <i>LightGbmRankingTrainer, FastTreeRankingTrainer</i> 															
<p>Entrenamiento del modelo</p>	<p>Este es el paso donde teniendo todo lo anterior se procede con el inicio del entrenamiento del modelo configurado con su respectivo algoritmo y con data inicialmente cargada. Par ello se hace uso de la función Fit (); la cual toma los datos, los aprende y crea un transformador el cual convierte los datos de entrada en las predicciones de salida que se requieren y da el modelo entrenado como salida.</p>															



SC-CER96940



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



Evaluación del modelo	<p>Para verificar la calidad que tiene el modelo de ML.NET, esta plataforma ofrece evaluadores que permiten diagnosticar el rendimiento y la precisión de las predicciones, entre ellos se destacan:</p> <ul style="list-style-type: none">• Precisión• Area bajo la curva o AUC• R2 o R-cuadrado• Error cuadrático medio o RMSE
Implementar y consumir modelo	<p>En este último paso y tener todo el modelo evaluado y con alto porcentaje de acierto, se puede realizar un guardado mediante archivo zip, lo cual permite el uso de este modelo en más aplicaciones .NET</p> <p>Para empezar a realizar las predicciones se hace uso de un motor de predicción que corresponde a una API en la cual se realizan únicas predicciones o por lotes haciendo uso de Transform.</p>

Con todo lo anterior, esta tecnología es muy amigable con el desarrollo web de las empresas actualmente, por el manejo de código abierto, ser multiplataforma, ser amigable con el lenguaje de programación de los expertos, contiene demasiadas opciones para la creación de modelos con alta eficiencia y confiabilidad, es por lo cual, a través de esta plataforma se realiza el presente proyecto para ser el inicio del uso de esta tecnología en la empresa Digital Ware, y permite seguir con la compatibilidad que llevan al usar tecnologías NET en los proyectos.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CAPITULO V. METODOLOGIA

Para la realización de este proyecto, se toma en cuenta la siguiente metodología, la cual consta de 4 fases que conforman la realización del clasificador de los errores de pruebas de automatización en la empresa de Digital Ware.

5.1 DESCRIPCION DEL LUGAR DE TRABAJO.

Digital Ware tiene la sede principal ubicada en la ciudad de Bogotá, pero con sedes secundarias en ciudades como Cali, Medellín, Bucaramanga y Barranquilla; y em países como Ecuador, Perú, México, Chile, Venezuela, Bolivia y Salvador, para la realización de este proyecto se hizo parte del equipo de automatización de la sede Bogotá, Calle 72 # 12 - 65 Piso 2, el cual es dirigido por el Ingeniero Edwin Barreto y conformado por 7 personas que realizan todo el proceso de automatización del producto Kactus-SCM, evaluando la calidad de los productos empresariales y reportando el proceso a área de calidad.

El tipo de metodología que se planteó para la realización del proyecto se basa en el modelo mixto, ya que se hace uso de datos cualitativos como datos cuantitativos, para el primer caso, se presenta en la fase 1 de la metodología planteada, ya que en ella se realizó un análisis característico de la información necesaria y adecuada para la aplicación del proyecto, tales como la elección del mejor modelo, técnica y algoritmo basándose en los atributos que presentan mejor acople al proceso a desarrollar, y también en la elección de las herramientas a usar para el proyecto, a través de las opiniones y valoraciones que se le dan de manera mundial a estas herramientas siendo las más precisas para el proyecto. En el segundo caso, para el análisis cuantitativo, se hace presente en las fases 2, 3 y 4 de la metodología, ya que en ellas se hace el análisis numérico de la información al ser adecuada en variables numéricas, se incursiona en la matemática del modelo, técnica y algoritmo escogido, con creaciones de algoritmos basados en tratamiento de variables enteras, y basándose en las métricas que dan validez al modelo creado.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



BASES TEÓRICO-PRÁCTICAS

- Selección y adquisición de las herramientas necesarias para la configuración del algoritmo.
- Configuración pruebas automatizadas y sus herramientas de manera correcta.

DATA SISTEMA

- Lanzamiento de las pruebas automatizadas.
- Recolección de la data necesaria para entrenar el algoritmo.
- Adecuar la información orientada al lenguaje de la adecuando para el algoritmo.

ENTRENAMIENTO

- Codificación del algoritmo con la data previamente adecuada, haciendo uso de ML.NET

VALIDACIÓN

- Validar que la clasificación de los errores se realice adecuadamente.

Ilustración 28. Metodología [Propia]

Como primera etapa de esta metodología planteada, se destacan dos secciones muy importantes, la primera es la selección de las herramientas necesarias para la creación del algoritmo, en esta sección, se hizo la selección del mejor modelo, técnica y algoritmo basándose en la investigación teórica realizada, teniendo aspectos importantes como la matemática del proceso, y en la segunda sección, se realiza la configuración de las pruebas automatizadas a través de herramientas de la empresa Digital Ware, como Visual Studio, Proyecto Web Kactus con las pruebas automatizadas previamente creadas con lenguaje C# y los pipelines en Azure, donde se configura la canalización y los agentes necesarios para el lanzamiento masivo de las pruebas automatizadas y obtener la mayor cantidad de resultados erróneos, la cual se convertirá en la data de entrenamiento.

Para la segunda etapa de la metodología, se realizan 3 secciones, la primera corresponde a inicial el lanzamiento de las pruebas automatizadas haciendo uso de la canalización en Azure en conexión con el proyecto web de la empresa, la cual facilita y agiliza el lanzamiento masivo de las pruebas y la visualización de los errores presentados para proceder con la extracción de los



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



mismos; en la segunda sección, se inicia con la revisión de cada uno de los resultados que dio el Pipeline en Azure, en las compilaciones resultantes, y la extracción de los textos de los errores presentados, para almacenarlos de manera momentánea en un archivo como Excel. Para luego en la última sección, adecuarlo de tal manera que sea editado en un archivo .txt donde se aplique el formato necesario para a carga de la data a través de ML.NET, es por ello que se acomoda de tal manera que la etiqueta y el texto del error sean separados por tabulaciones y la librería al hacer el llamado sepa diferenciar la posición de las columnas.

Para la tercera etapa de la metodología, se inicia con la ubicación de la data adecuada en el directorio correspondiente a la creación de la aplicación de consola con NET. CORE, y se prosigue con sumar al proyecto las librerías ML.NET agregando Nuggets a la solución para que sea importada de manera general en todo el código, el código se divide en 4 partes como lo son, la caracterización y carga de la información, la creación del modelo, la evaluación del modelo y finalmente la predicción que realiza basándose en el texto aleatorio que se ingresa para mirar su efectividad y comportamiento.

Finalmente se da la etapa de validación donde teniendo todo el modelo, se realiza el ingreso de errores aleatorios resultantes de pruebas automatizadas para validar con varios intentos la calidad y predicción del modelo, a su vez analizando la medición de esta calidad con el uso de las métricas estipuladas para la técnica de clasificación binaria.

5.2 DESARROLLO DE LA METODOLOGIA

FASE 1

1.1 Selección y adquisición de selección y adquisición de las herramientas necesarias para la configuración del algoritmo.

Inicialmente se llevó a cabo el proceso de elección de las herramientas que permitirán la estructuración del modelo de clasificación de errores, las cuales se destacan en el siguiente gráfico.



SC-CER96940

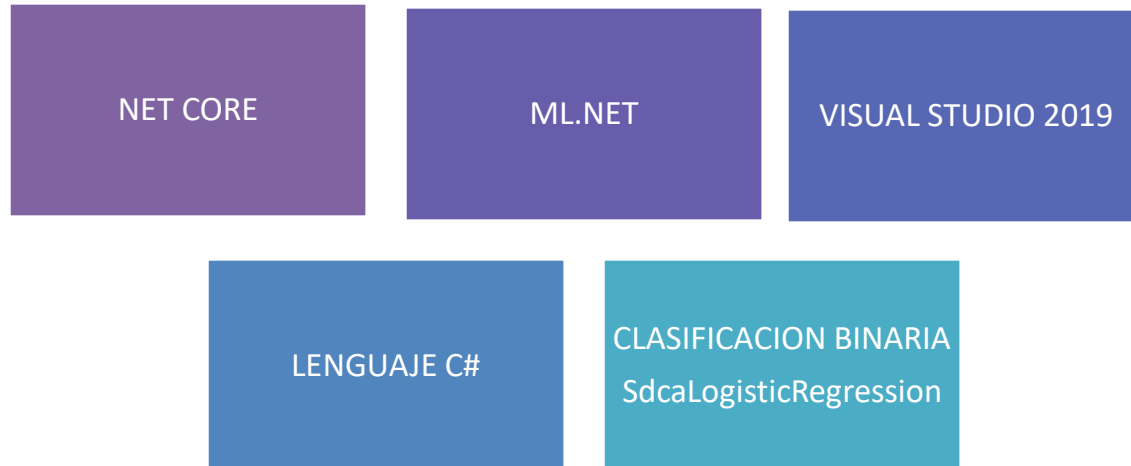


“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Como primera herramienta se destaca NET CORE, esta plataforma es una sección de .NET que permite la creación de aplicaciones de consolas en la cual se realizara la ejecución del proyecto de manera correcta debido al manejo que tiene con el lenguaje C#, la facilidad de usar más tecnologías de .NET sin tener problemas de compatibilidad. Es un ambiente agradable para la creación del modelo de predicción de errores.

La segunda herramienta es la más importante de todas, ML.NET va a permitir la codificación del código personalizado para la creación del modelo de machine Learning, esta plataforma al tener como base la tecnología NET.CORE permite su ejecución de manera correcta, cuenta como muchas librerías y compatibilidades otras bibliotecas como Tensorflow y ONNX, dando aún más facilidades en la creación del algoritmo, con ML.NET se puede realizar la carga de los datos, creación del contexto donde se va a ejecutar el código, la canalización de conexión, entrenar el modelo, evaluarlo y finalmente comprobar su precisión.

La tercera herramienta corresponde a el ambiente de trabajo que mejor se comportaba en el uso de las dos tecnologías mencionadas anteriormente de manera paralela, para la creación de la aplicación de NET. CORE, la administración del NuGet de ML.NET, la facilidad de interactuar con demás tecnologías de .NET, Visual Studio 2019 es el mejor software para este proyecto, mantiene muchas actualizaciones de tecnologías al día permitiendo trabajar con tecnologías de ultimo lanzamientos.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!

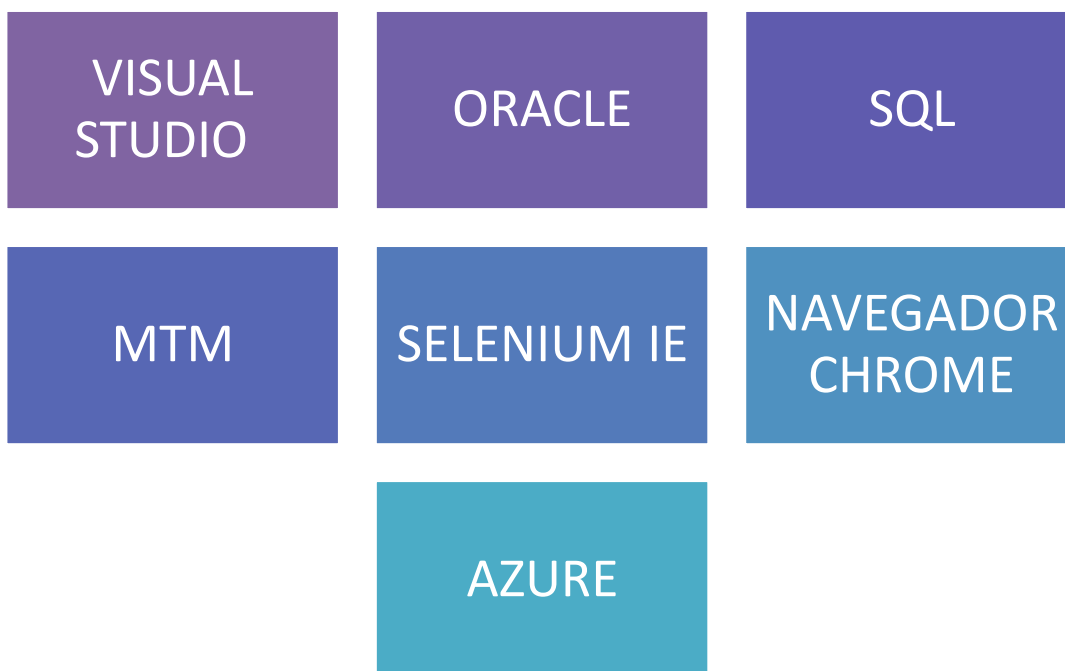


El lenguaje de programación escogido es el C#, es el lenguaje que mejor trabaja las aplicaciones de ML.NET actualmente permite el fácil entendimiento del código con el creador, y a su vez es el lenguaje que la empresa usa a nivel de programación de pruebas automatizadas, por lo cual genera una compatibilidad acorde para hacer uso de las mismas, y extracción de los datos de manera sencilla y práctica.

Finalmente, se eligió el tipo de machine Learning que se va a desarrollar además de su modelo y su técnica, en el ámbito del modelo, el proyecto se basa en clasificar errores basándose en etiquetas por lo cual corresponde al aprendizaje supervisado en la sección de clasificación, aquí se destacan dos clases la binaria y la multiclase, como en el proyecto desarrollado se determinaba la predicción entre dos estados, un error por programador y otro por sistema, entonces se realizó la elección de clasificación binaria. Teniendo lo anterior finalmente se hizo la elección de la técnica, en este tipo de clasificación se destacaban varios algoritmos de alto nivel, pero el que mejor se acomodó al sistema fue el SdcaLogisticRegression, el cual se podía usar en la tecnología ML.NET.

1.2 Configuración pruebas automatizadas y sus herramientas de manera correcta.

Para la configuración de las pruebas automatizadas a utilizar para la extracción de data y para la validación del funcionamiento, se hace necesario el uso de las herramientas:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Primera herramienta necesaria para las pruebas automatizadas, hace referencia a Visual Studio, este es el ambiente donde actualmente el equipo de automatización de Digital Ware realiza las creaciones de los códigos, es una herramienta que permite la escritura en lenguajes de programación variados, pero a nivel empresarial se usa el C#, por lo cual facilita la compatibilidad con la creación del modelo de machine Learning planteado.

La segunda y tercera herramienta son las bases de datos que la compañía usa actualmente para el manejo de la información que se usan para las pruebas, también son herramientas que permiten encolar las pruebas que se desean ejecutar haciendo uso de las características que nos ofrece el manejo de los casos con la herramienta Microsoft Test Manager.

La cuarta herramienta hace referencia al Microsoft Test Manager (MTM), es una plataforma que permite la conexión con el servidor principal de la empresa donde se crean las variables necesarias para el almacenamiento de los datos para ser llamados por el código y realizar la prueba, tales como el usuario, contraseña, url, entre otros. En esta aplicación se almacenan los datos generales de cada prueba para ser encolada en las tablas de las bases de datos.

Las dos herramientas, Selenium IDE y Chrome, son las correspondientes al ambiente donde se hace la interacción de las pruebas con la aplicación de la empresa, las pruebas del proyecto web Kactus, son automatizadas en su totalidad en el navegador de Chrome al ser el más estable en mantener la ejecución de las pruebas automatizadas. Selenium IDE es una herramienta importante que permite la lectura de los xpath de los campos, variables, botones, títulos entre otros componentes de la página, los cuales son de mucha importancia en la creación del código para cada una de las pruebas. Para el uso de esta herramienta se hace a través de las extensiones del navegador Chrome donde se ingresa el sitio a interactuar y se da click en los elementos que se necesitan y este devuelve el xpath del valor del botón, campo o texto a nivel de lenguaje de página.

Finalmente se presenta la herramienta de Azure, en esta herramienta se realiza la configuración de pipelines encargados de activar la ejecución de pruebas en máquinas específicas ubicadas en la empresa, donde se realiza la configuración de un agente, se activan las opciones para actualizar las tablas de encolamiento de manera automática, y el lanzamiento masivo de las pruebas basándose en los datos guardados en el MTM. A través de Azure se realiza el lanzamiento pruebas automatizadas.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



FASE 2

2.1 Lanzamiento de pruebas automatizadas.

Para realizar este proceso, se tomó como base las pruebas automatizadas del proyecto web las cuales corresponden de alrededor 400 pruebas tanto para el motor de Oracle como para SQL, y así lograr la mayor cantidad de data necesaria para el entrenamiento del algoritmo, para ello se usarán las pruebas dadas a continuación:

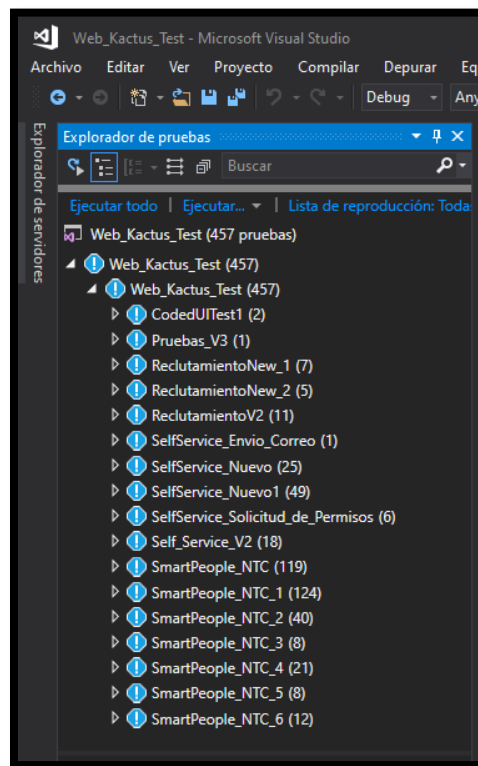


Ilustración 29. Pruebas automatizadas [Propia]

Como se observa, en la ilustración anterior, estas son las pruebas automatizadas presentes actualmente en el proyecto Web Kactus, las cuales serán la fuente principal para la extracción de los errores necesarios para el entrenamiento del algoritmo, donde cada módulo contiene una serie de métodos donde se realiza la configuración de la prueba en un programa específico. De manera general se manejan dos estilos de pruebas como lo son Reclutamiento y Self Service, estas son pruebas continuas que validan el correcto funcionamiento de los registros y aprobaciones de solicitudes, las pruebas Smart People NTC son pruebas de validación donde se



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



prueba si el estado de la página web esta de manera correcta, como campos, nombres botones etc. A continuación, un ejemplo del código de estas pruebas, de cómo es la estructura y el esquema que se maneja de manera interna. Son programadas en lenguaje C# en la herramienta Visual Studio, con uso mayoritario de librerías Selenium, debido al trato de automatizaciones web, creando un repositorio que se conecta con el servidor Azure de la empresa donde se realiza el lanzamiento masivo a la par de lo que se desarrolla. A continuación, se muestra un ejemplo de una prueba codificada en las herramientas empresariales.

```
11538 }
11539 }
11540 }
11541 }
11542 [TestMethod]
11543 [Diferencias | cambios | autores - cambios
11544 public void SmartPeople_AdicionDeUVTNTC()
11545 {
11546     List<string> errorMessages = new List<string>();
11547     List<string> errorMessagesMetodo = new List<string>();
11548     bool bandera = false;
11549     string enviroment = (Environment.MachineName);
11550     string[] auxatable = enviroment.Split('-');
11551     string TableOrder = "";
11552     if (auxatable.Length > 1)
11553     {
11554         TableOrder = (enviroment.Replace("-", "_")).ToUpper();
11555     }
11556     else
11557     {
11558         TableOrder = enviroment.ToUpper();
11559     }
11560     // TableOrder = "ktes1";
11561     DataSet OrderExecutionCase = SqlDataAdapter.SelectOrderExecution("", TableOrder);
11562     int NumCasAgen = OrderExecutionCase.Tables[0].Rows.Count;
11563     if (NumCasAgen < 1)
11564     {
11565         errorMessages.Add("No hay casos en el agendamiento");
11566     }
11567     foreach (DataRow row in OrderExecutionCase.Tables[0].Rows)
11568     {
11569         string plans = row["plans"].ToString();
11570         string suite = row["suite"].ToString();
11571     }
11572 }
```

Ilustración 30. Ejemplo código pruebas automatizadas [Propia]

Para le correcto funcionamiento de estas pruebas, se hace necesario el uso del MTM donde se almacenan las variables con sus datos necesarias para la ejecución correcta del código, seria de la siguiente forma, este MTM es obligatorio para el control de las variables que el código necesita para su funcionamiento, y administración de la información por parte de mi como programadora.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Contenido

Nuevo Agregar requisitos

SelfService

- NuevosCasosSelfService
 - Nuevo conjunto
 - NuevosSelfServiceORA
 - NuevosSelfServiceSQL
 - PruebasReqPersonalSQL
 - PruebasGeneralesORACLE
 - PruebasGeneralesSQL
 - PRUEBASORA
 - PruebasReqPersonalMIempresaORACLE
 - PruebasReqPersonalMIempresaSQL
 - PruebasReqPersonalORACLE
 - PRUEBASSQL
 - SalvarORACLE
 - SalvarSQL
 - SelfService_EnvioCorreo_ORA
 - SelfService_EnvioCorreo_SQL
 - SmartPeople_NTC_ORA_1
 - SmartPeople_NTC_ORA_10
 - SmartPeople_NTC_ORA_11
 - SmartPeople_NTC_ORA_12
 - SmartPeople_NTC_ORA_2
 - SmartPeople_NTC_ORA_3
 - SmartPeople_NTC_ORA_4
 - SmartPeople_NTC_ORA_5
 - SmartPeople_NTC_ORA_6
 - SmartPeople_NTC_ORA_7
 - SmartPeople_NTC_ORA_8
 - SmartPeople_NTC_ORA_1

Conjunto de pruebas: SmartPeople_NTC_ORA_1 (Id. de conjunto: 1869)
Configuraciones predeterminadas (1): Windows 8

Abrir Agregar Nuevo X Asignar Configuraciones Ordenar

Arrastre aquí un encabezado de columna para agrupar por esa columna.

Orden	Id	Título	Priori...	Conf...	Evaluaadores	Ruta de acceso del área
1	1872	SmartPeople_ReporteManualFuncionesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
2	1873	SmartPeople_SolicitudinscripcionRPCINTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
3	1899	SmartPeople_OrganigramapresidenciaNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
4	1900	SmartPeople_CarguePersonalAuditoriasNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
5	1901	SmartPeople_FrmBelncopNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
6	1902	SmartPeople_ReportefinancieroNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
7	1903	SmartPeople_NecesidadesDeformacionNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
8	1904	SmartPeople_frmBeSolicColINTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
9	2064	SmartPeople_MisIdiomasNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
10	2066	SmartPeople_MisDatosBasicosAunNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
11	2067	SmartPeople_ExamenesMedicosNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
12	2068	SmartPeople_AdministracionActivosNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
13	2069	SmartPeople_ListadoRolesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
14	2070	SmartPeople_RolesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
15	2071	SmartPeople_PolizasNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
16	2072	SmartPeople_ProyectosLogrosNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
17	2073	SmartPeople_MisPublicacionesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
18	2074	SmartPeople_ServicioAlClienteInternoNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
19	2075	SmartPeople_BeneficiosOrganizacionalesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
20	2076	SmartPeople_MisCentrosVacacionalesNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
21	2077	SmartPeople_ListaColaboradoresNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
22	2078	SmartPeople_MisPrestamosNTC_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
23	2079	SmartPeople_EntrevistaDeRetiro_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus
24	2080	SmartPeople_MisSolicitudesCertificaciones_ORA	2	1	Kactus SCM < DIGIT...	WEB Kactus

Ilustración 31. Datos almacenados en MTM [Propia]

Caso de prueba 1872: SmartPeople_ReporteManualFuncionesNTC_ORA

SmartPeople_ReporteManualFuncionesNTC_ORA 1872

Iteración WEB Kactus

ESTADO
Asignado a Kactus SCM
Estado Diseño
Área WEB Kactus

DETALLES
Prioridad 2
Estado de automatización Automatizada

PASOS RESUMEN ELEMENTOS DE TRABAJO PENDIENTE PRUBADOS VÍNCULOS DATOS ADJUNTOS AUTOMATIZACIÓN ASOCIADA

Insertar paso Insertar pasos compartidos Insertar parámetro

Acción	Resultado esperado
1. Login	@url @EmpleadoUser @EmpleadoPass
2. Variables	@TituloEsp @SubtituloEsp @HomeEsp @GuardarEsp @RegresarEsp @ConCodCargEsp @ConNomCargEsp @CodCargEsp @NomCargEsp @VersionEsp @VigenciaEsp @FechaElabEsp @GrupoEsp @CodigoEsp @EstadoEsp @ElaboradoEsp @AprobadoEsp @ValidadoEsp @AutorizadoEsp @DescripcionEsp @ObservacionesEsp

Haga clic aquí para agregar un paso

Valores de parámetro

Eliminar iteración Cambiar el nombre del parámetro Eliminar parámetro

url	EmpleadoUser	EmpleadoPass	TituloEsp	SubtituloEsp	HomeEsp	GuardarEsp	RegresarEsp	ConCodCargEsp	ConNomCargEsp	CodCargEsp	NomCargEsp	VersionEsp	VigenciaEsp	FechaElabEsp
http://dvtfskcm/se...	79907078	1234	Manual de Descripci...	Manual de Descripci...	Inicio	Guardar	Regresar	Consultar codigo car...	Consultar nombre ca...	Codigo cargo	Nombre cargo	Versión	Vigencia desde	* Fecha elabora...

Ilustración 32. Caso en MTM [Propia]



“Formando líderes para la construcción de un nuevo país en paz”
 Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
 www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



En la ilustración 31 se observa cómo se clasifican los casos de acuerdo las normas estéticas de la empresa como lo son máximo 40 casos por suite, entre otras. Mientras que en la ilustración número 32 se observa cómo se almacenan las variables necesarias para el buen funcionamiento de las pruebas tales como el nombre, el contenido, numero caso y nombre caso etc.

Teniendo el MTM y el código listo, se procede a realizar las configuraciones en Azure, para la creación de los pipelines donde almacenaran los módulos de las pruebas automatizadas a lanzar, para ello se realizó el siguiente procedimiento:

- Primero se hace la creación del pipeline, para ello se muestra el proceso con uno y ese mismo se repite para todos los pipelines que correspondan a las pruebas automatizadas a lanzar. Para el ejemplo se creará un pipeline para las pruebas correspondientes a los módulos NTC_1 Y NTC_2 para el motor de Oracle.

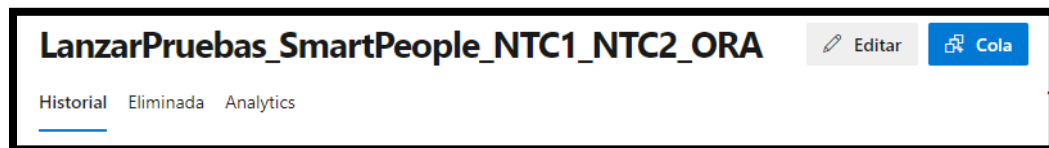


Ilustración 33. Creación Pipeline [Propia]

En la configuración general del proyecto se debe verificar que el agente que tiene el equipo de la empresa se encuentre activo, así como lo muestra la siguiente imagen, es muy importante para que se realice la conexión correcta entre maquina y pipeline. Este agente por políticas y administración de la empresa, viene configurado a nivel de CMD en cada una de las maquinas presentes en el dominio de Digital Ware, para su activación en la configuración de pipeline, como se muestra a continuación:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Ilustración 34. Grupo Agentes [Propia]

- Teniendo ya creado el pipeline, se da click en **Editar**, desde allí se abren varias secciones necesarias:
 - Canalización: En esta sección se configura el grupo de agentes encargado de mantener la conexión para que el lanzamiento se realice de manera correcta, en este caso se configuro con el nombre de la maquina a usar DW-A1143.

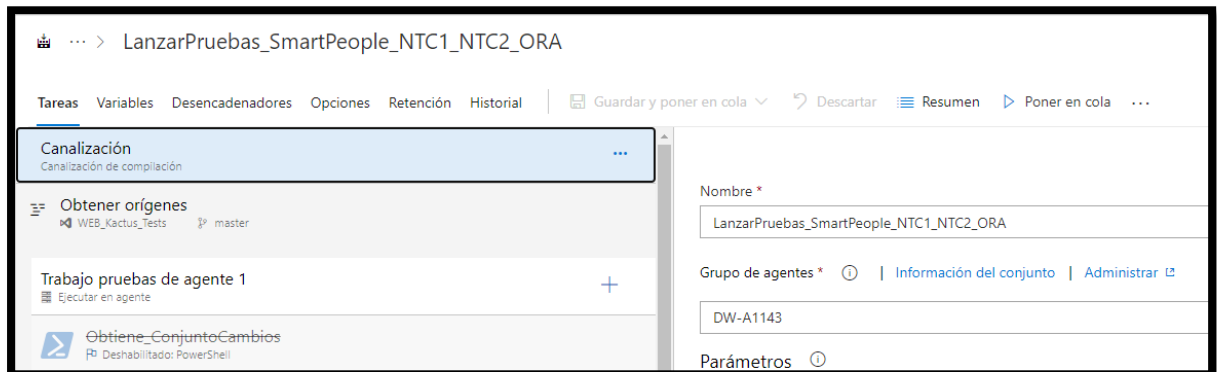


Ilustración 35. Canalización [Propia]

- Siguiendo con el proceso de configuración, se hace la creación de los siguientes parámetros.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

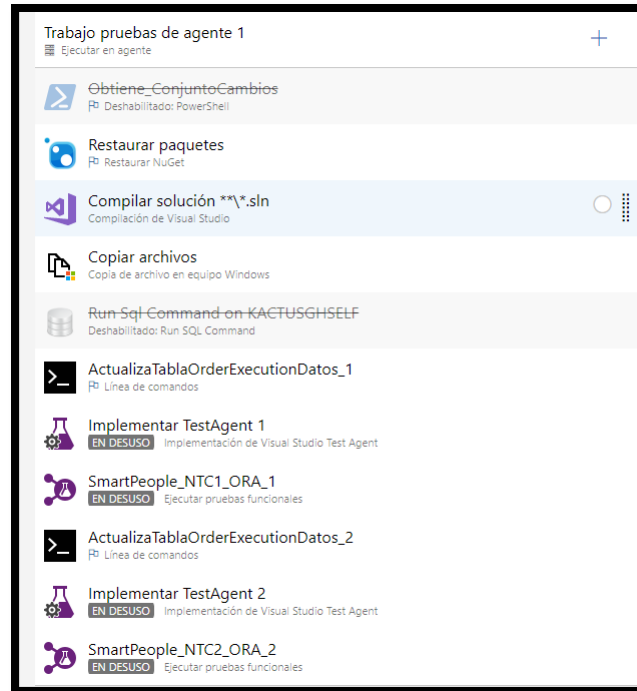


Ilustración 36. Parámetros pipeline [Propia]

- Restaurar Paquetes: Es una sección donde se hace la descarga de todos los NuGets necesarios para el correcto desarrollo del código.
- Compilar Solución: Es la sección donde se realiza la configuración de las herramientas donde se desarrolló la codificación de las pruebas y a su vez compilarla para ratificar su correcto funcionamiento, aquí se configura dónde está la solución, en que programa se desarrolla, VS 2017, plataforma y configuración.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *

Compilar solución ***.sln

Solución * ⓘ

***.sln

Versión de Visual Studio ⓘ

Visual Studio 2017

Argumentos de MSBuild ⓘ

/p:OutDir="\$(build.binariesdirectory)\\"

Plataforma ⓘ

\$(buildPlatform)

Configuración ⓘ

\$(buildConfiguration)

Limpiar ⓘ

Ilustración 37. Compilar Solución [Propia]

- Copiar archivos: Sección donde se facilita al equipo los archivos necesarios para la recepción de las pruebas por parte del pipeline.
- ActualizaTablaOrderExecutionDatos_1: Aquí se facilita la ruta específica donde se almacena el programa de TargetCase que es un programa encargado de actualizar la tabla de encolamiento con la información de las pruebas a lanzar y el orden en el que se lanzarán, este programa debe estar ubicado en la máquina donde se lanzan las pruebas, es decir, DW-A1143. También se añade el nombre de las pruebas a ejecutar, el id del pipeline, la máquina donde se lanzarán y demás características mostradas a continuación:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *	ActualizaTablaOrderExecutionDatos_1
Herramienta * ⓘ	E:\TfsScripts\TargetCase\TargeCases.exe
Argumentos ⓘ	"SmartPeople_NTC1_ORA_1" "60" "\${Maquinas}" "TEAMSERVER"
Avanzado ▾	
Opciones de control ▾	
Variables de salida ▾	

Ilustración 38. Actualizar Tabla [Propia]

- Implementar TestAgent: Se facilita la información más importante necesaria para activar el agente y crear la conexión con la maquina a ejecutar las pruebas, se le anexa en que maquina está el agente, con que usuario y clave se accede con permisos de administrados, al igual que el usuario y contraseña de la sesión, todo esto es manejado con variables globales para mantener la información en total confidencialidad.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *

Máquinas de prueba ^

Máquinas * ⓘ

Inicio de sesión del administrador * ⓘ

Contraseña de administrador * ⓘ

Protocolo * ⓘ
 Http Https

Configuración de agente ^

Nombre de usuario * ⓘ

Contraseña * ⓘ

Ejecutar pruebas de IU ⓘ

Ilustración 39. TestAgent [Propia]

- SmartPeople_NTC_1_ORA: Es la sección donde se especifica el grupo de pruebas almacenadas en la herramienta MTM que se van a lanzar, primero se nombra de la misma manera que se configuro en la sección de implementación de agente, luego se especifica que plan se va a usar en este caso solo existe **Test Plan**, que sección de este plan se va a ejecutar, caso **Self Service**, y finalmente el conjunto de pruebas que se van a lanzar, estas opciones son sacadas de la misma manera que están organizadas las pruebas en el MTM.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *
SmartPeople_NTC1_ORA_1

Opciones de instalación ^

Máquinas * ⓘ
\$(Maquinas)

Ubicación de almacenamiento de la prueba * ⓘ
\$(testAssemblyDropLocation)

Opciones de ejecución ^

Selección de pruebas * ⓘ
Test Plan

Plan de pruebas * ⓘ
SelfService

Conjunto de pruebas * ⓘ
SelfService\NuevosCasosSelfService\SmartPeople_NTC_ORA_1

Ilustración 40. Pruebas a lanzar [Propia]

- Este mismo proceso se realiza con el módulo NTC_2, configurando nuevamente el agente, actualizar la tabla de encolamiento con los nuevos datos y finalmente se crea otra sección, donde se cambia el conjunto de pruebas por el que se está configurando.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *
ActualizaTablaOrderExecutionDatos_2
Herramienta * ⓘ
E:\TfsScripts\TargetCase\TargeCases.exe
Argumentos ⓘ
"SmartPeople_NTC2_ORA_2" "60" "\$(Maquinas)" "TEAMSERVER"
Avanzado ▾
Opciones de control ▾
Variables de salida ▾

Ilustración 41. Actualizar Tabla 2 [Propia]

Se ingresa los datos adecuados para el las pruebas del módulo NTC_2, donde se ingresa la misma información que el anterior lo único que varía es los nombres.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *
Implementar TestAgent 2

Máquinas de prueba ^

Máquinas * ⓘ
\$(Maquinas)

Inicio de sesión del administrador * ⓘ
\$(User)

Contraseña de administrador * ⓘ
\$(Pass)

Protocolo * ⓘ
 Http Https

Configuración de agente ^

Nombre de usuario * ⓘ
\$(User)

Contraseña * ⓘ
\$(Pass)

Ejecutar pruebas de IU ⓘ

Ilustración 42. Test Agent 2 [Propia]

En la anterior ilustración se muestra cómo se configura el segundo agente para el lanzamiento de NTC_2 donde se usa la misma información que la primera configuración.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Nombre para mostrar *	SmartPeople_NTC2_ORA_2
Opciones de instalación ^	
Máquinas * ⓘ	\$(Maquinas)
Ubicación de almacenamiento de la prueba * ⓘ	\$(testAssemblyDropLocation)
Opciones de ejecución ^	
Selección de pruebas * ⓘ	Test Plan
Plan de pruebas * ⓘ	SelfService
Conjunto de pruebas * ⓘ	SelfService\NuevosCasosSelfService\SmartPeople_NTC_ORA_2

Ilustración 43. Conjunto Pruebas 2 [Propia]

En la ilustración 43 se muestra la selección del conjunto de pruebas correspondientes al módulo NTC_2 y acomodando el nombre para diferenciarlo de la primera configuración, teniendo así todo definido para dar por finalizada la configuración del pipeline en su totalidad.

- Finalizada las configuraciones de las secciones necesarias se guarda el proceso y se **Pone en cola** el pipeline para que empiece a realizar la preparación e inicio de las pruebas automatizadas, como se muestra a continuación:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



Variables	Demandas
system.debug	false

Ilustración 44. Poner en cola [Propia]

Confirmación	N.º de compilación
KS 10112021 Compilación manual para Kactus SCM	14364

Ilustración 45. Compilaciones del Pipeline [Propia]

Todo el proceso realizado anteriormente se desarrolla para los demás módulos presentes en las pruebas automatizadas, dando como resultado la siguiente lista de pipelines a la espera de lanzamiento para obtener los resultados de las pruebas automatizadas.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
LanzarPruebas_NuevosSelfService_PruebasGenerales_ORA
LanzarPruebas_NuevosSelfService_PruebasGenerales_SQL
LanzarPruebas_Reclutamiento_V2ORACLE
LanzarPruebas_Reclutamiento_V2SQL
LanzarPruebas_SelfService_RequisicionesORA
LanzarPruebas_SelfService_RequisicionesSQL
LanzarPruebas_SelfService_V2ORACLE
LanzarPruebas_SelfService_V3SQL
LanzarPruebas_SmartPeople_EnvioCorreo_ORA
LanzarPruebas_SmartPeople_EnvioCorreo_SQL
LanzarPruebas_SmartPeople_NTC1_NTC2_ORA
LanzarPruebas_SmartPeople_NTC1_NTC2_SQL
LanzarPruebas_SmartPeople_NTC11_NTC12_ORA
LanzarPruebas_SmartPeople_NTC11_NTC12_SQL
LanzarPruebas_SmartPeople_NTC3_NTC4_ORA
LanzarPruebas_SmartPeople_NTC3_NTC4_SQL
LanzarPruebas_SmartPeople_NTC5_NTC6_ORA
LanzarPruebas_SmartPeople_NTC5_NTC6_SQL
LanzarPruebas_SmartPeople_NTC7_NTC8_NTC10_ORA
LanzarPruebas_SmartPeople_NTC7_NTC8_NTC10_SQL
```

Ilustración 46. Pipelines [Propia]

Y se realiza el lanzamiento masivo de todos estos pipelines de acuerdo a la disponibilidad de la maquina ubicada en Digital Ware, esta ejecución continua de pruebas es de larga duración por lo cual se distribuyó en horas diarias su compilación para evitar saturaciones y sobreponer ejecuciones.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



2.2 Recolección de la data necesaria para entrenar el algoritmo.

Para realizar esta sección de la fase 2, se procede a ingresar al Azure y observar en cada uno de los pipelines creados anteriormente, accediendo a la compilación realizada y observar los resultados erróneos y así extraer la mayor cantidad de errores presentados y almacenarlos en un Excel como borrador, para luego realizar su respectiva adecuación.

- Se accede a Azure, sección de Pipelines y observar el proceso realizado por la compilación obtenida de cada uno de los lanzamientos de los pipelines con módulos de pruebas automatizadas, y en la sección de **Pruebas**, se observa el porcentaje de las pruebas correctas e incorrectas, y lo más importante para la extracción de la data, los errores que se presentaron. En la siguiente imagen, se observa cómo se muestran los resultados al acceder a la compilación resultante, mostrando el paso a paso que se configuro en cada uno de los pipelines.

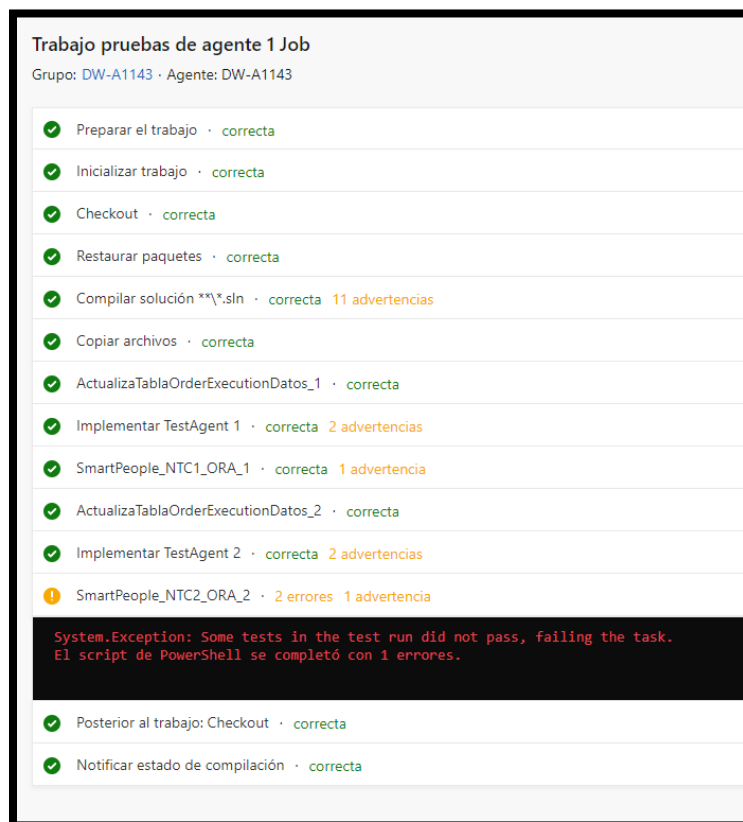


Ilustración 47. Proceso compilación desarrollado [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

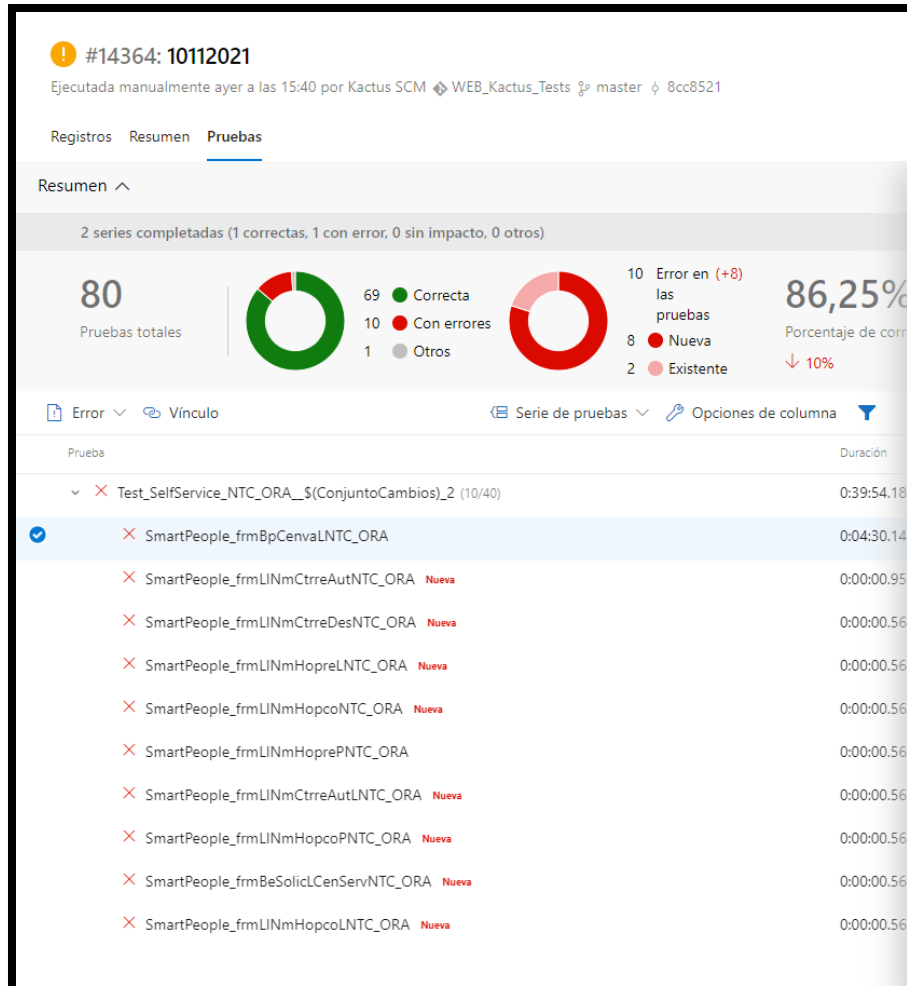


Ilustración 48. Resultados Pruebas [Propia]

Al dar click en **Pruebas**, se observa todas las características globales de los resultados haciendo un balance de las correctas y las que presentaron errores, representándose como se da en la ilustración 48.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



Para obtener la data, es necesario extraer todos los errores posibles, para ello se ingresa a cada una las pruebas con errores y se extrae el mensaje de error que muestran, como por ejemplo el siguiente:

```
SmartPeople_frmBpCenvaLNTC_ORA
X Con errores hace 14 horas en DW-A1143 Duración 0:04:30.144
Compilación con errores 14341 Propietario Servicio de compilación de colecciones ...
Depurar Elementos de trabajo Datos adjuntos Historial
Mensaje de error ^
Error de Assert.Fail. 1921 :::::OpenQA.Selenium.WebDriverException: The HTTP request to the remote
http://localhost:54159/session/a4616314fc7edd082b5a50236f214ba4/element timed out after 240 seconds
tiempo de espera de la operación
en System.Net.HttpWebRequest.GetResponse()
en OpenQA.Selenium.Remote.HttpCommandExecutor.MakeHttpRequest(HttpRequestInfo requestInfo)
--- Fin del seguimiento de la pila de la excepción interna ---
en OpenQA.Selenium.Remote.HttpCommandExecutor.MakeHttpRequest(HttpRequestInfo requestInfo)
en OpenQA.Selenium.Remote.HttpCommandExecutor.Execute(Command commandToExecute)
en OpenQA.Selenium.Remote.DriverServiceCommandExecutor.Execute(Command commandToExecute)
en OpenQA.Selenium.Remote.RemoteWebDriver.Execute(String driverCommandToExecute, Dictionary`2 param
en OpenQA.Selenium.Remote.RemoteWebDriver.FindElement(String mechanism, String value)
en OpenQA.Selenium.Remote.RemoteWebDriver.FindElementByXPath(String xpath)
en OpenQA.Selenium.By.<>c__DisplayClass19_0.<XPath>b__0(ISearchContext context)
en OpenQA.Selenium.By.FindElement(ISearchContext context)
en OpenQA.Selenium.Remote.RemoteWebDriver.FindElement(By by)
en APITest.APISelenium.Click(String control) en C:\TfsAgentDWA1143\_work\1\s\SeleniumAPITest\APISel
en Web_Kactus_Test.SmartPeople_NTC_1.SmartPeople_frmBpCenvaLNTC() en C:\TfsAgentDWA1143\_work\1\s\W
Seguimiento de la pila ^
en Web_Kactus_Test.SmartPeople_NTC_1.SmartPeople_frmBpCenvaLNTC() en C:\TfsAgentDWA1143\_work\1\
```

Ilustración 49. Errores Presentados [Propia]

Finalmente se extrae todos estos mensajes de errores en un borrador, en este caso, se hizo uso de la herramienta Excel para llevar una mejor organización del proceso, como se muestra a continuación:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



A	B	C	D	E	F
etiqueta	errores				
1	OpenQA.Selenium.ElementNotInteractableException: element not interactable			1 sistema	
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Lugar Administrativo y el encontrado es: Lugar administrativo			0 nosotros	
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Tipo de Operación y el encontrado es: Tipo de				
1	Message: Test 'SmartPeople_frmBIDataadNTC' exceeded execution timeout period.				
1	element click intercepted: Element <span id="ctl00_ContentoPagina_KCtrITxtValMTEX_FORT_lblTexto"				
1	: The HTTP request to the remote WebDriver server for URL http://localhost:65306/session/36dc89627b554ac8a422655a6cc0ed97/element timed out after 240 seconds. ...				
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Identificación Jefe y el encontrado es: Identificación				
0	MSG: El Nombre Emergente es incorrecto, el esperado es: N. Contrato y el encontrado es: Nro. Contrato				
0	MSG: El título es incorrecto, el esperado es: Mis Solicitudes de Paz y Salvos con Flujo de Aprobación y el				
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Nro. Carnet y el encontrado es: Nro.carnet				
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Hora de Inicio HH/MM y el encontrado es: Hora de				
0	MSG: El Nombre Emergente es incorrecto, el esperado es: Nombres y Apellidos y el encontrado es: Nombres				
0	El caso de prueba presento los siguientes errores:				
0	El caso de prueba presento los siguientes errores:				
0	El caso de prueba presento los siguientes errores:				
1	The HTTP request to the remote WebDriver server for URL http://localhost:15041/session/ef5ac89110ed91521839fa2c618855be/element timed out after 240 seconds				
0	El caso de prueba presento los siguientes errores:MSG: El Nombre Emergente es incorrecto, el esperado es: N.				
0	El caso de prueba presento los siguientes errores:				
1	The HTTP request to the remote WebDriver server for URL http://localhost:15560/session/42a5ff2c217af4e6f3ff74bf7eb2d444/element timed out after 240 seconds				
1	The HTTP request to the remote WebDriver server for URL http://localhost:15665/session/d57b96fc49d12486b0c3abf7491e7ddf/element timed out after 240 seconds				
1	The HTTP request to the remote WebDriver server for URL http://localhost:15893/session/1d84d61a71ccee259c739b938824374/element timed out after 240 seconds				
1	Referencia a objeto no establecida como instancia de un objeto.				
1	The HTTP request to the remote WebDriver server for URL http://localhost:16134/session/290cf10a043d69184174409f94d66470/element timed out after 240 seconds,				
0	El caso de prueba presento los siguientes errores:				
0	El caso de prueba presento los siguientes errores:				
0	El caso de prueba presento los siguientes errores:				
1	element click intercepted: Element <input name="ctl00\$ContenidoPagina\$StxtNomEmpl" type="text"				

Ilustración 50. Errores obtenidos de los resultados pruebas automatizadas [Propia]

2.3 Adecuar la información orientada al lenguaje de la adecuando para el algoritmo.

Con el uso de la plataforma ML.NET, para que el algoritmo detecte cuales son las filas y columnas correctas, se debe manejar la mejor estructura para llevar un orden de los datos, se deben hacer modificaciones en el tipo de archivo, ML.NET acepta de manera correcta los archivos donde:

- Formato CSV Delimitado por comas de Excel.
- Delimitado por tabulaciones en formato txt.
- Valores separados por tabulaciones en formato tsv de Excel.

Para los datos recolectados, se hizo la adecuación de la información en un archivo .txt en la cual se coloca primero el error y con tabulaciones se separa de la columna de etiqueta, quedando de la siguiente manera:



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
data.txt: Bloc de notas
Archivo Edición Formato Ver Ayuda
element not interactable Session info chrome. 1
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Lugar Administrativo y el encontrado es Lugar administrativo.
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Tipo de Operación y el encontrado es Tipo de operación. 0
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Tipo de Operación y el encontrado es Tipo de operación. 0
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Centro de Costos y el encontrado es Centro de costos. 0
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Valor Solicitado y el encontrado es Valor solicitado. 0
Message Test exceeded execution timeout period. 1
element click intercepted Element is not clickable at point (635, 732). 1
The HTTP request to the remote WebDriver server for URL element timed out after 240 seconds Anulada la solicitud Se excedió el tiempo de espera de la operación.
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Identificación Jefe y el encontrado es Identificación jefe.
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Nombre Jefe y el encontrado es Nombre jefe. 0
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Identificación Autorizador y el encontrado es Identificación au
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Nombres Autorizador y el encontrado es Nombres autorizador.
El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es N. Contrato y el encontrado es Nro. Contrato. 0
El caso de prueba presento los siguientes errores: MSG: El título es incorrecto, el esperado es: Mis Solicitudes de Paz y Salvos con Flujo de Aprobación y el encontrado
El caso de prueba presento los siguientes errores: MSG: El Nombre Emergente es incorrecto, el esperado es: Nro. Carnet y el encontrado es: Nro.carnet MSG: El Nombre Em
El caso de prueba presento los siguientes errores: MSG: El Nombre Emergente es incorrecto, el esperado es: Hora de Inicio HH/MM y el encontrado es: Hora de inicio HH/MM
El caso de prueba presento los siguientes errores: MSG: El Nombre Emergente es incorrecto, el esperado es: Consultar por apellidos y el encontrado es: Consultar por
El caso de prueba presento los siguientes errores: MSG: El Nombre Emergente es incorrecto, el esperado es: Sub Proceso y el encontrado es: Sub proceso. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Tipo Vinculacion y el encontrado es: Tipo vinculacion. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Estado Cargo y el encontrado es: Estado cargo. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Inicial y el encontrado es: Fecha inicial. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Final y el encontrado es: Fecha final. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Inicial y el encontrado es: Fecha inicial. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Final y el encontrado es: Fecha final. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Inicial y el encontrado es: Fecha inicial. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha Final y el encontrado es: Fecha final. 0
El caso de prueba presento los siguientes errores: MSG: El título es incorrecto, el esperado es: Resumen de la Solicitud Aprobada y el encontrado es: Resumen de la sol
MSG: El subtítulo es incorrecto, el esperado es: Resumen de la Solicitud Aprobada y el encontrado es: Resumen de la solicitud aprobada. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: N. de Solicitud y el encontrado es: N. de solicitud. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Estado Actualizado y el encontrado es: Estado actualizado. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Persona que Actualiza y el encontrado es: Persona que actualiza 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Fecha de Actualizacion y el encontrado es: Fecha de actualización. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Nombre Solicitante y el encontrado es: Nombre solicitante. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Identificación Solicitante y el encontrado es: Identificación solicitante. 0
MSG: El Nombre Emergente es incorrecto, el esperado es: Cargo Solicitante y el encontrado es: Cargo solicitante. 0
```

Ilustración 51. Datos adecuados formato txt con tabulaciones [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



FASE 3

3.1 Inicio de codificación del algoritmo con la data previamente adecuada, haciendo uso de ML.NET

- Como primer paso para el desarrollo del algoritmo en ML.NET es abrir un nuevo proyecto en Visual Studio.

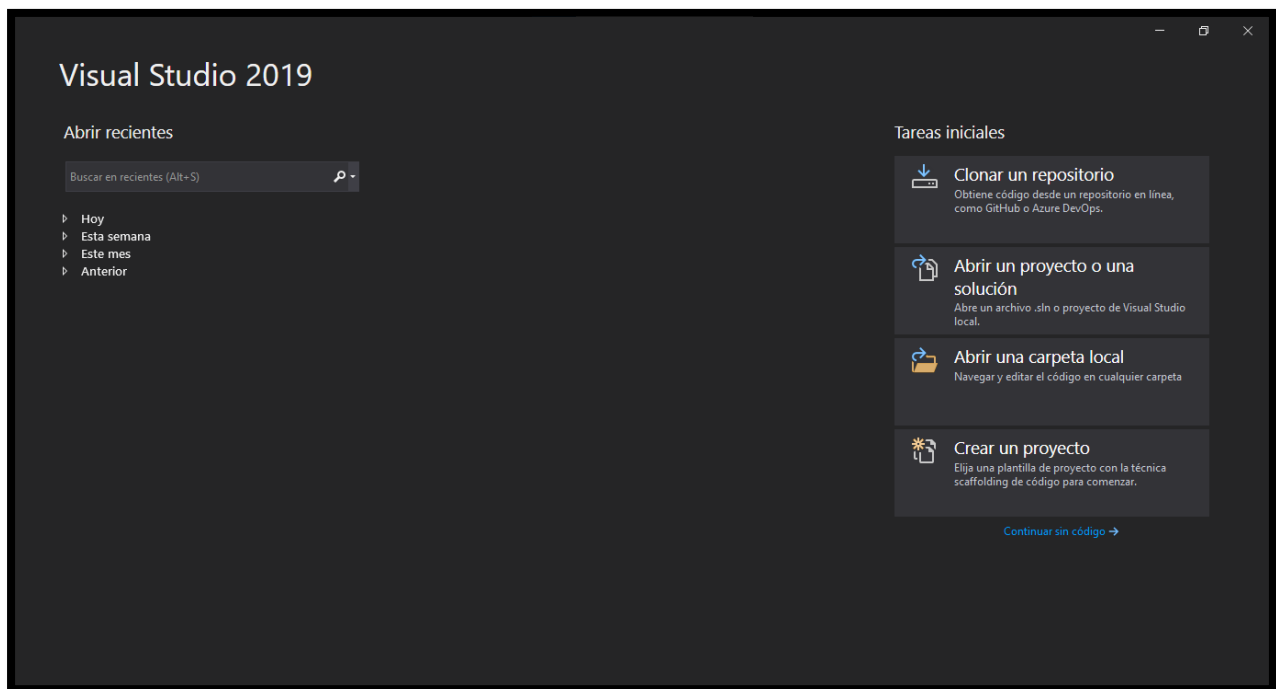


Ilustración 52. Visual Studio [Propia]

- Se busca la aplicación de consola, que corresponde al escenario de NET. Core haciendo uso del campo de búsqueda.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

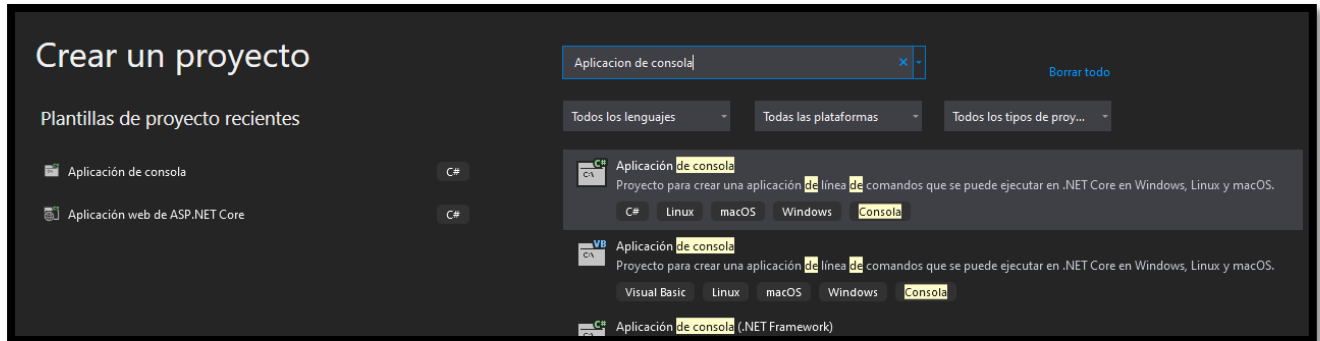


Ilustración 53. Aplicación de consola [Propia]

- Se añade el nombre del proyecto y también el nombre de la solución, en este caso se hizo uso del que se muestra a continuación.

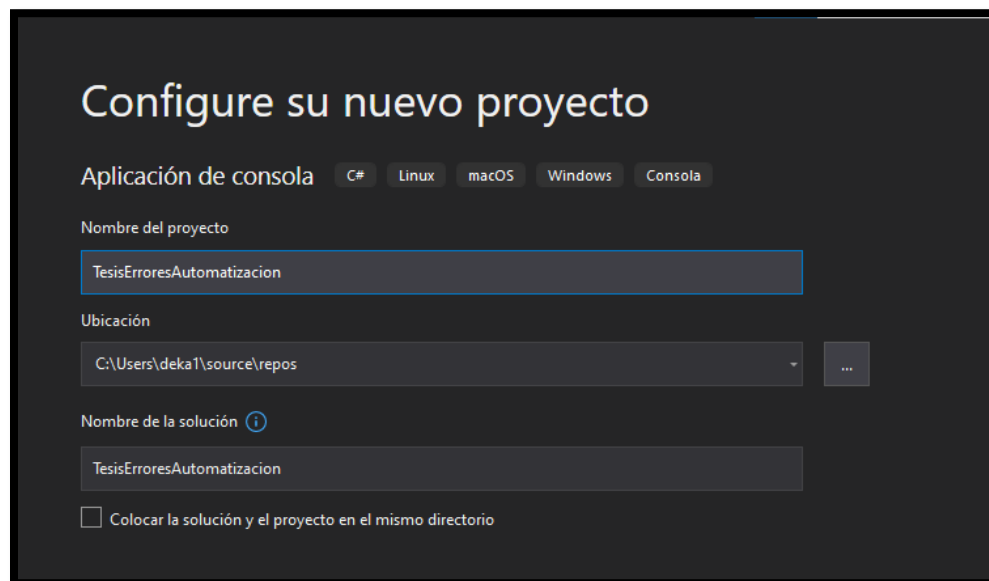


Ilustración 54. Nombre Proyecto y Solución [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Se agrega la versión de Net. Core, para compatibilidad con futuras actualizaciones, la mejor versión es la 3.1 que tiene un desarrollo compatible con aplicaciones futuras.

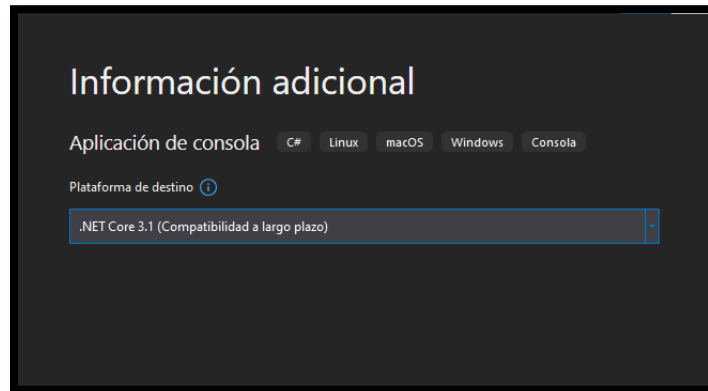


Ilustración 55. Versión Net. Core [Propia]

- Se da inicio a la creación de la aplicación y se obtiene el siguiente ambiente de trabajo por defecto.

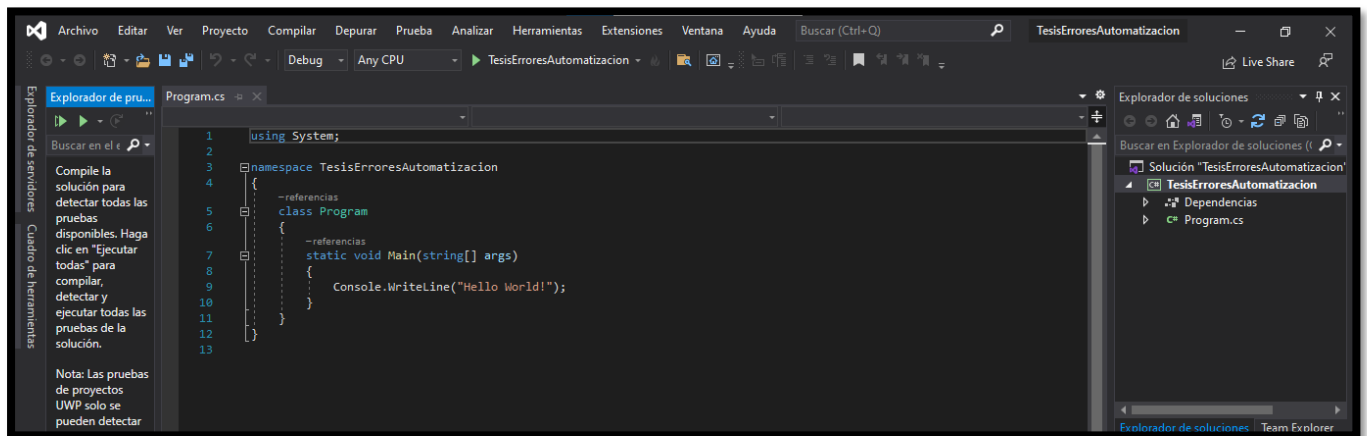


Ilustración 56. Ambiente de trabajo [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Ahora teniendo ya creada la aplicación de consola, se prosigue a agregar la plataforma ML.NET adjuntándolo al proyecto a través del administrador de NuGets, para ello se va al explorador de soluciones y se da click derecho en la solución.

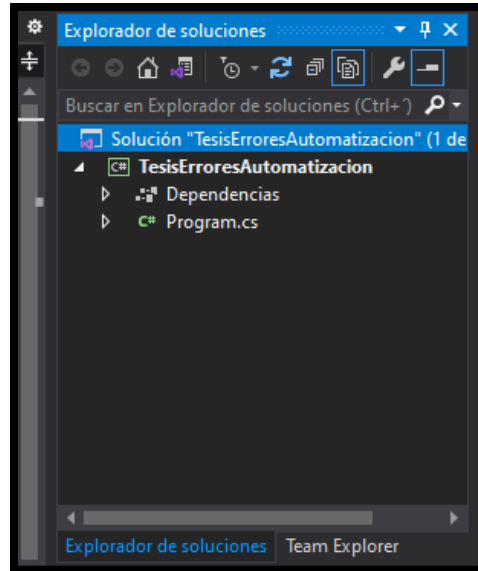


Ilustración 57. Solución Proyecto [Propia]

- Luego se despliega un menú, en donde se busca la opción **Administrar paquetes NuGets para la solución**, y se da click en ella

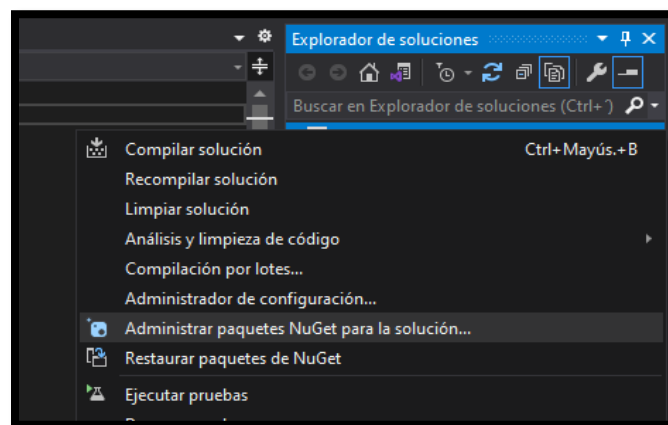


Ilustración 58. Administrador de NuGets [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Se abre un menú, en el cual se hace la búsqueda del NuGet llamado **Microsoft.ML**, se selecciona la primera opción, se realiza el check para que se aplique a todo el proyecto y se inicia el proceso de instalación, al presentarse la licencia de los productos se acepta y se sigue con el proceso.

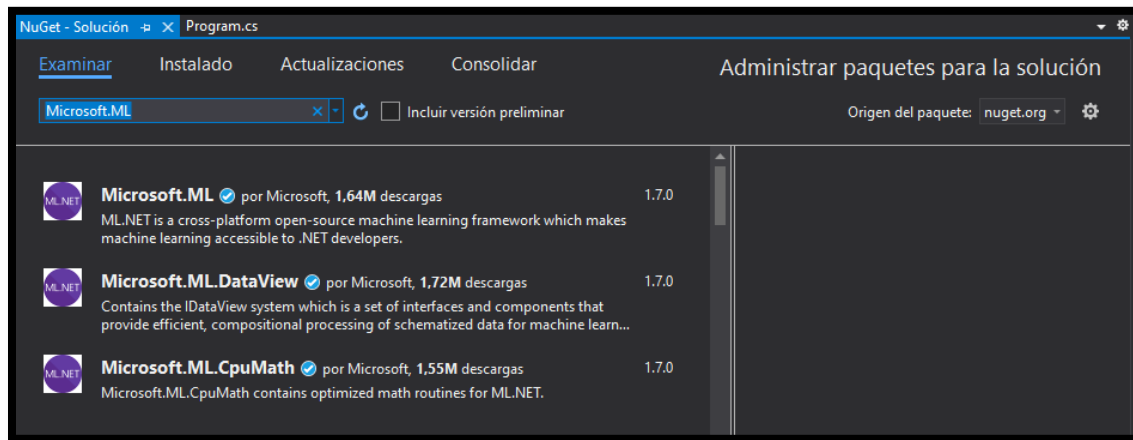


Ilustración 59. NuGet ML [Propia]

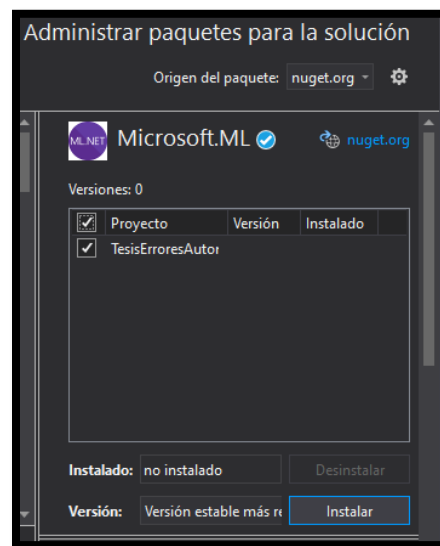


Ilustración 60. Instalación NuGet ML [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Ahora, se debe agregar el archivo txt con la data adecuada en la fase anterior en la ubicación de la solución, para ello se da click derecho a la solución, se busca la opción **abrir carpeta en el explorador de archivos**

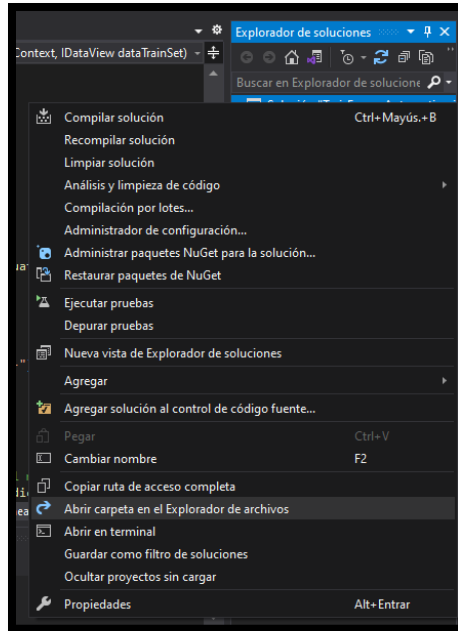


Ilustración 61. Abrir carpeta en Explorador de archivos [Propia]

- Se abre el explorador de archivos de Windows, y se accede a la siguiente ruta

```
C:\Users\deka1\source\repos\TesisErroresAutomatizacion\TesisErroresAutomatizacion\bin\Debug\netcoreapp3.1
```

- Cuando ya se esté dentro de la carpeta Debug/netcoreapp3.1, se inicia el proceso de la creación de un directorio con el nombre **Data** y allí se copiará el archivo txt con la data que se usara para el entrenamiento del algoritmo llamado **data.txt**



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

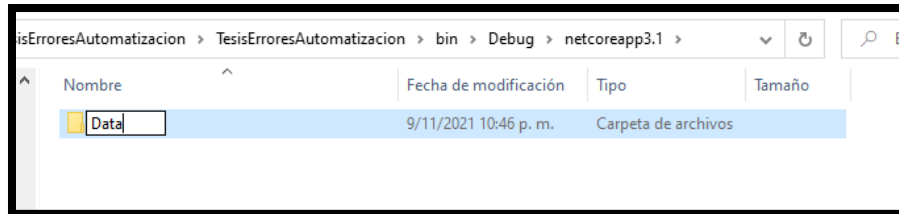


Ilustración 62. Directorio Data [Propia]

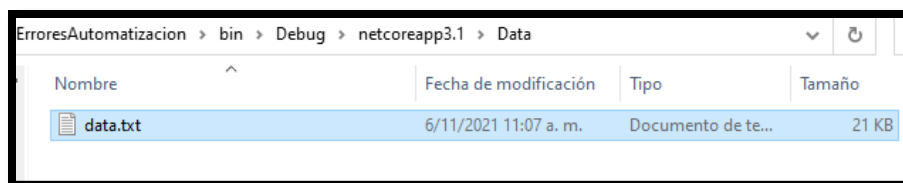


Ilustración 63. Copia data adecuada en txt. [Propia]

- Teniendo ya lista la data a usar en el directorio correspondiente y el NuGet de ML, se inicia con la creación de la primera clase necesaria para el inicio de la programación del algoritmo, para ello se va a aplicación, en el explorador de soluciones y dar click derecho

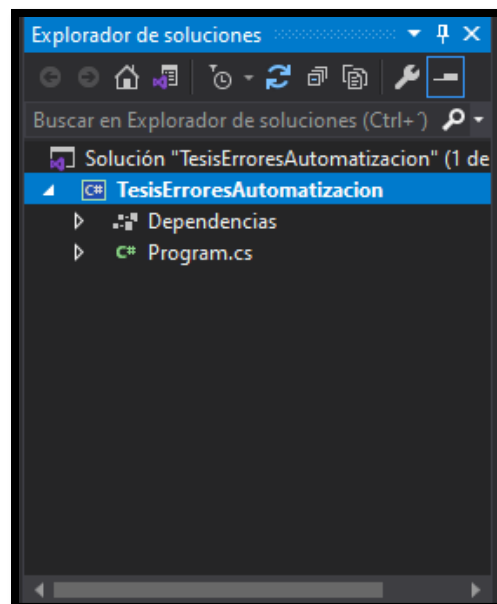


Ilustración 64. Aplicación en explorador de soluciones [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Se busca la opción de agregar y se selecciona la opción de **Nuevo Elemento**

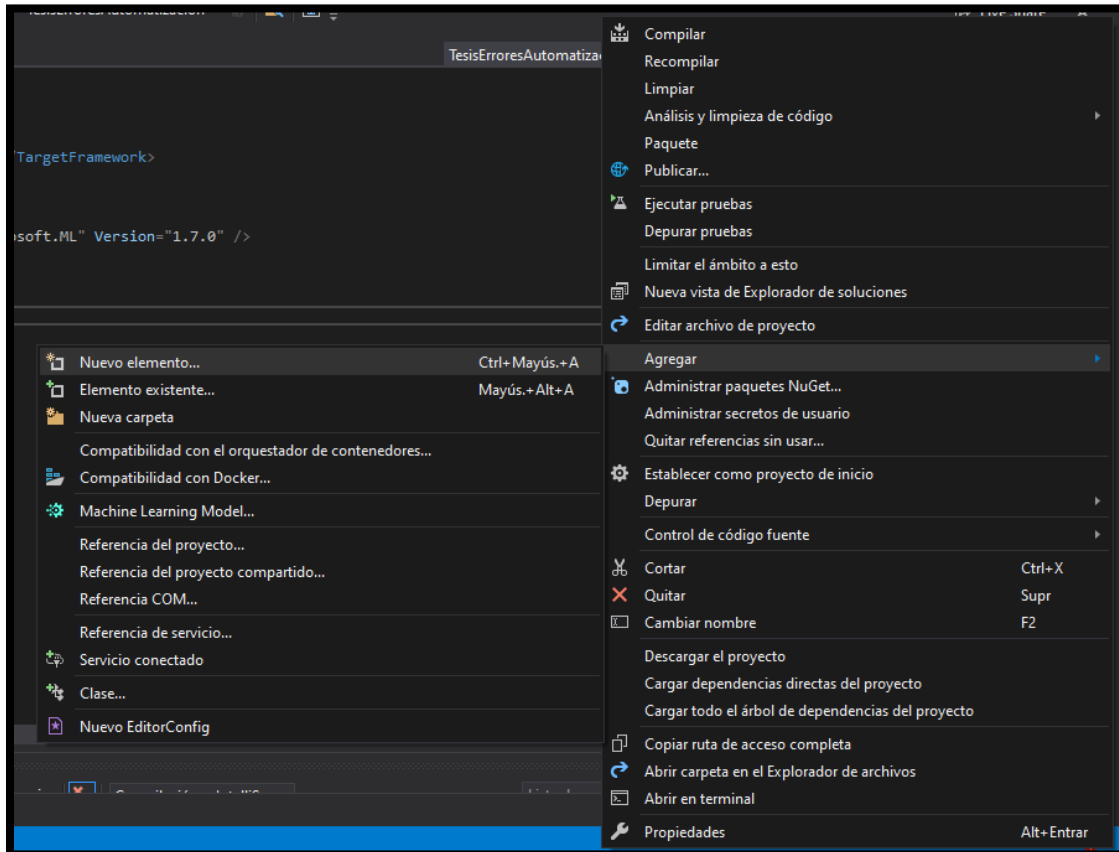


Ilustración 65. Agregar nuevo elemento [Propia]

- Luego se escoge el archivo de **clase** y se coloca el nombre **MistakeData**, en esta clase se realizará todo lo correspondiente a las especificaciones de la carga de la data.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

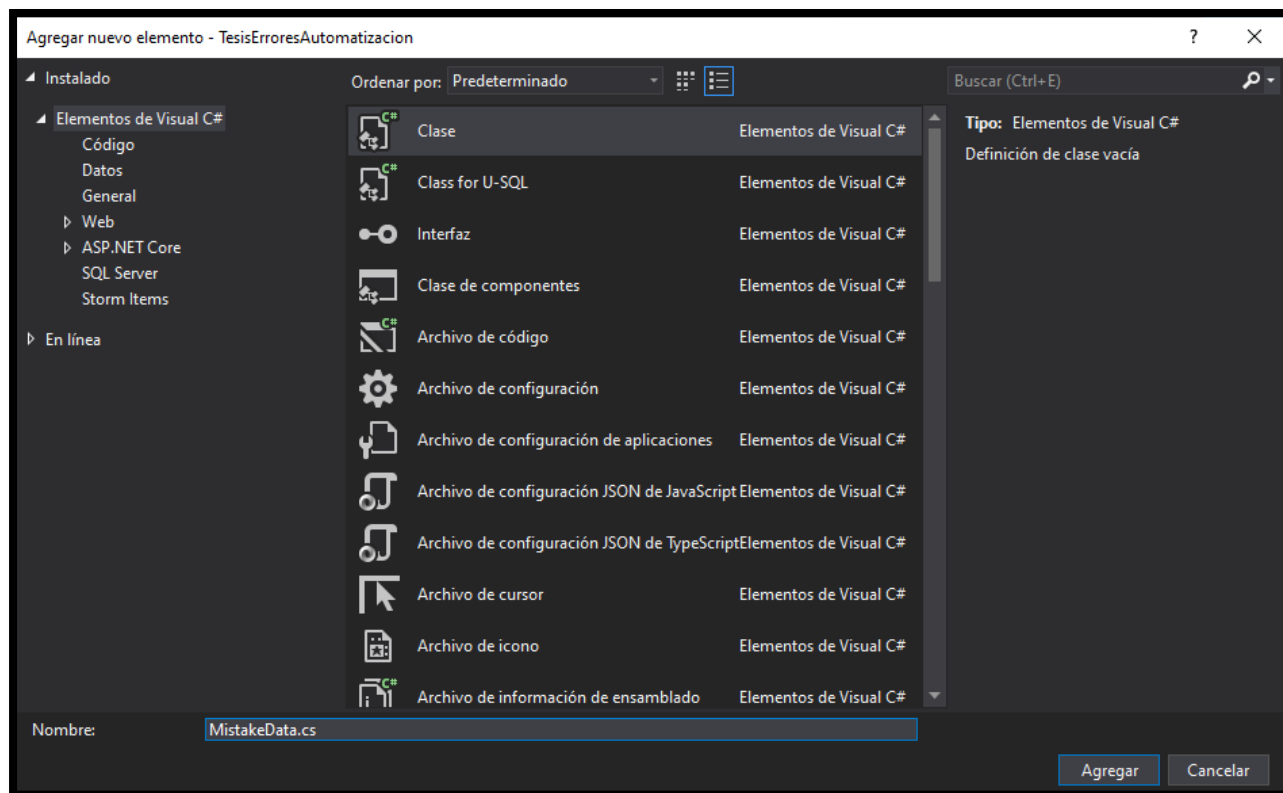


Ilustración 66. Clase MistakeData [Propio]

- Al tener ya esta clase creada, se inicia con la programación. Como primer paso se hace el llamado de las librerías a usar para esta clase, haciendo uso del using. La Librería a usar se llama

```
using Microsoft.ML.Data;
```

```
using System;  
using System.Collections.Generic;  
using System.Text;  
using Microsoft.ML.Data;
```

Ilustración 67. Librerías clase MistakeData [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Luego dentro de la clase de MistakeData, se procede a hacer la carga de las columnas presentes en la data de entrenamiento y que contiene cada una de ellas:
 - LoadColumn---- corresponde a la carga de la columna del archivo
 - 0 y 1---- hacen referencia al orden en el que están las columnas
 - Public string MistakeText---- le aclara al código que la información en la columna cero contiene los textos de los errores.
 - Public bool Mistake---- le aclara al código que la información que se encuentra ahí hace referencial estado del error, si es por sistema o programador.
 - ColumnName("Label") ---- nombra que el contenido en la columna número 1 de los datos es considerada como la etiqueta y así sepa identificar el atributo de clasificación.

```
namespace TesisErroresAutomatizacion
{
    7 referencias
    public class MistakeData
    {
        // Carga de los datos haciendo uso de las columnas donde se encuentran y cual es la etiqueta e informacion
        [LoadColumn(0)]
        public string MistakeText;

        [LoadColumn(1), ColumnName("Label")]
        public bool Mistake;
    }
}
```

Ilustración 68. Código Clase MistakeData [Propia]

- Después de clase MistakeData se crea una nueva llamada **MistakePrediction** que hereda información de la función anterior, en donde almacenara:
 - [ColumnName("PredictLabel")]---- Asigna el nombre de PredictLabel y una columna para que almacene la información de las variables que almacenan la predicción, probabilidad y puntaje de la respuesta del sistema.
 - public bool Prediction {get; set;} ---- Variable de predicción
 - public float Probability {get; set;} ---- Variable de probabilidad
 - public float Score {get; set;} ---- Variable de puntaje

Las tres últimas variables corresponden a las métricas de evaluación para sistemas que manejan clasificación binaria donde `Score` hace relación al puntaje que se comportara como umbral como se explicó en la sección teórica, para clasificación binaria. `Prediction` es la variable donde se almacena la respuesta que da el algoritmo de



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



acuerdo a su entrenamiento y `Probability` corresponde a la variable que almacena la probabilidad resultante del algoritmo basándose en regresión logística.

Estas variables obtienen data y dan información se ahí sus propiedades `get` y `set`, dando paso a ejecución de varias predicciones.

```
2 referencias
public class MistakePrediction : MistakeData
{
    //Se adjuntan las 3 variables de las cuales se basa la prediccion como lo es probabilidad, produccion y puntaje
    [ColumnName("PredictLabel")]
    1 referencia
    public bool Prediction { get; set; }

    1 referencia
    public float Probability { get; set; }

    0 referencias
    public float Score { get; set; }
}
```

Ilustración 69. Clase MistakePrediction [Propia]

- Con estas dos funciones se da por finalizada la programación en la clase de `MistakeData` para el sistema en desarrollo. Ahora se sigue la programación en la clase **Program**, la cual venia por defecto en la creación de la consola, estando ahí se hace el llamado de varias librerías que se usaran al avanzar con el código, estas serían:

```
using System;
using System.Collections.Generic;
using System.IO;
using System.Linq;
using Microsoft.ML;
using Microsoft.ML.Data;
using static Microsoft.ML.DataOperationsCatalog;
using Microsoft.ML.Trainers;
using Microsoft.ML.Transforms.Text;
```



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
using System;  
using System.Collections.Generic;  
using System.IO;  
using System.Linq;  
using Microsoft.ML;  
using Microsoft.ML.Data;  
using static Microsoft.ML.DataOperationsCatalog;  
using Microsoft.ML.Trainers;  
using Microsoft.ML.Transforms.Text;
```

Ilustración 70. Librerías clase Program [Propia]

- Estando en la clase **Program**, como primer paso, dentro del método main que viene por defecto, se crea el contexto en el cual se hará el desarrollo del algoritmo ya que en ese espacio se desarrolla todo el entrenamiento y validación.

```
static void Main(string[] args)  
{  
    // crea el contexto necesario para la configuracion del algoritmo.  
    MLContext mlContext = new MLContext();  
}
```

Ilustración 71. Contexto ML.NET [Propia]

- Luego al comenzar la clase Program y afuera de la función Main, se realiza la programación de una línea de código en la cual se explica de manera general la ubicación exacta en la que se encuentra el archivo txt con la data necesaria.
 - Readonly ----solo lectura
 - Path----permite realizar acciones en el directorio o archivo que contiene la información
 - Combine ---- permite que se especifique la ruta haciendo uso de tres datos path.
 - Environment.CurrentDirectory ----esta sección permite obtener todas las características del directorio donde se trabajará, seguidamente dando el nombre de la carpeta donde se encuentra y como se llama el archivo con información de entrenamiento.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
class Program
{
    //crea una variable de solo lectura, la cual tiene la ruta específica del archivo.
    static readonly string load = Path.Combine(Environment.CurrentDirectory, "Data", "data.txt");
    Referencias
    static void Main(string[] args)
    {
    }
}
```

Ilustración 72, Ubicación archivo txt [Propia]

- Fuera de la clase Program, se crea el método llamado **LoadData**, en él se hace la configuración y llamado de los datos para usar en el entrenamiento.
 - IDataView--- permite describir los datos que tengan formato de tabulación y contengan datos numéricos y de texto.
 - dataView--- es la variable donde se almacenan los datos.
 - mlContext---siempre debe hacerse una acción basándose en el contexto creado en la función main.
 - Data --- librería de ml.net que permite la administración de los datos.
 - LoadFromTextFile--- permite definir el esquema que maneja los datos y la lectura de los archivos. Para ello se le deben pasar la función donde se definió que contiene cada columna y cuál es la etiqueta, en este caso **MistakeData**, seguido se da la ruta del archivo que se declaró al comienzo de la clase el cual se llamó **load**, y el ultimo parámetro a pasar es si el documento donde se almacena la información tiene una primera fila tiene encabezado o no, en este caso no la tiene por eso se hace el envío de **hasHeader: false**.
 - TrainTestData--- elemento que permite el manejo de los datos de tal manera que una parte pueda ser para entreno y otra para evaluar el algoritmo.
 - Data---variable que guarda la cantidad de datos que se usan para el entreno
 - TrainTestSplit--- es un método de ML.NET que permite dividir los datos cargados en datos de entrenamiento y datos de prueba, para ello se le pasan dos parámetros importantes, donde están los datos cargados, es decir **dataView**, y la fracción que se quiere para entrenamiento, se escogió el 50% del total, haciendo uso de **testFraction:0.5**.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
//CARGA DE DATA
1 referencia
public static TrainTestData LoadData(MLContext mlContext)
{
    //se hace la carga de los datos haciendo uso de loadfromtextfile
    IDataView dataView = mlContext.Data.LoadFromTextFile<MistakeData>(load, hasHeader: false);

    //especifica el porcentaje de datos totales se van a usar para el entrenamiento, se escoge el 50%
    TrainTestData data = mlContext.Data.TrainTestSplit(dataView, testFraction: 0.5);

    //retorna la variable con los datos que se van a usar para el entrenamiento del modelo
    return data;
}
```

Ilustración 73. LoadData [Propia]

- Después del método LoadData, se crea el método donde se construye y entrena el modelo, el cual llamaremos **Model**, el cual va a recibir dos parámetros, el contexto de ML y los datos que se asignaron para el entrenamiento.
 - Este método se basa en `ITransformer`, un componente que permite la transformación de los datos, para mejor entendimiento del algoritmo.
 - `Transform`— variable que almacena los datos de entrenamiento transformados
 - `Transform.Text.FeaturizeText` --- funcionalidad de `ml.net` la cual permite transformar los datos de entrenamiento en el lenguaje natural de la máquina, en este caso, se dan dos parámetros, **outputColumnName : "Features"**, en la cual se almacenaran todos los datos de clave numérica resultado de la transformación agregándola al conjunto de datos que usa para el entrenamiento del modelo, **inputColumnName : nameof(MistakeData.MistakeText)**, hace referencia a la columna de entrada donde se almacenan los errores.
 - `Append`--- es una funcionalidad de `ML.net` la cual permite agregar un parámetro a la línea que se lleva, en este caso se agregara con que algoritmo se va a realizar el modelo de clasificación, para así crear el pipeline del modelo.
 - `BinaryClassification`—al ser una clasificación de solo dos estados en las predicciones se hace el llamado de este parámetro que permite el total manejo de problemas con esta característica.
 - `Trainers`--- sección dentro del catálogo de clasificación binaria donde se encuentran los entrenadores de algoritmos.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- o `SdcaLogisticRegression` --- es la técnica dentro de los entrenadores que mejor se comporta en problemas de clasificación binaria, donde hace uso de la regresión lineal para realizar la correcta predicción del tipo de error, para su funcionamiento se envían dos parámetros importantes como lo es la columna de los datos de entrenamiento que corresponde a la etiqueta, el cual llamamos en la clase `MistakeData` como **Label**, y el segundo parámetro es la columna de características que se desarrolló en el paso anterior donde se hizo la transformación de los errores de texto a clave numérico, el cual llamamos **Features**.
- o `Model`--- es la variable donde se va a guardar el modelo entrenado
- o `Transform`--- se hace el llamado de la transformación de texto realizada al comienzo del método.
- o `Fit`--- es una función de `ML.net` que permite iniciar el entrenamiento del modelo teniendo en cuenta los datos asignados para el mismo, el cual se llamó **dataTrainSet**, y devuelve el modelo ya finalizado y entrenado, con las especificaciones realizadas respecto al algoritmo escogido.

```
// CONSTRUCCION DEL MODELO
1 referencia
public static ITransformer Model(MLContext mlContext, IDataView dataTrainSet)
{
    //Realiza la transformación de los datos de tal manera que sea entendible por la maquina.
    var transform = mlContext.Transforms.Text.FeaturizeText(outputColumnName: "Features", inputColumnName: nameof(MistakeData.MistakeText))

    // Elige el algoritmo que mejor describa la clasificación que se desea de errores.
    .Append(mlContext.BinaryClassification.Trainers.SdcaLogisticRegression(labelColumnName: "Label", featureColumnName: "Features"));

    Console.WriteLine("-----Inicio de construccion del modelo-----");

    //Crea el modelo haciendo uso de la función Fit
    var model = transform.Fit(dataTrainSet);

    Console.WriteLine("-----Finalizacion de construccion del modelo-----");

    //Finalmente retorna el modelo listo para su evaluacion
    return model;
}
```

Ilustración 74. Model [Propia]

- Teniendo ya el modelo listo, se procede a crear un método para evaluar el modelo haciendo uso de métricas, disponibles gracias a `ML.Net`, este se va a llamar **Evaluate**, el cual va a recibir tres parámetros, el contexto `ML`, el modelo entrenado y el porcentaje de datos fraccionados para evaluar, el cual corresponde al otro 50% de la data cargada.



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- `IDataView`--- se vuelve a usar este escenario ya que se va a ser uso de la fracción de data asignada para evaluar el modelo.
- `dataTest` --- donde se realiza la transformación de la data de evaluación del modelo, de tal manera que sea entendible para el modelo, por ello se hace uso del mismo.
- `CalibrateBinaryClassificationMetrics`--- es una función de ML.net que permite evaluar las mediciones que se realizan en clasificadores binarios.
- `Measurement` --- variable donde se almacena la evaluación que se hace al modelo de clasificación binaria basándose en la función `Evaluate`.
- `Evaluate`--- Evalúa los puntajes de la clasificación binaria de la data asignada para evaluación, es por eso que se le deben ingresar dos parámetros, la data transformada que se usara para evaluar, la cual se llamó **dataTest**, y el segundo parámetro es el nombre de la columna donde se almacena la etiqueta que clasifica los errores.
- Después se hace uso de 3 métricas de evaluación que se usan en modelos de clasificación binaria, como lo son `Accuracy` (Precisión): permite conocer el porcentaje de precisión del modelo entrenado basándose en datos de evaluación, entre más cercano este de 100% mejor es la calidad del modelo. `AreaUnderRocCurve` (Area bajo la curva), permite evaluar la tasa de predicciones verdaderamente positivas y falsas positivos, si esta entre el 50 % y 100% es considerado un modelo aceptable. Y finalmente `F1Score` es una métrica exclusiva de clasificación binaria, que abarca dos valores que dan que tan preciso es el clasificador, 0 si es pesimo y 1 si es muy bueno.

```
//EVALUACION DEL MODELO
1 referencia
public static void Evaluate(MLContext mlContext, ITransformer model, IDataView dataTestSet)
{
    Console.WriteLine("-----Evaluacion del modelo-----");
    //Transforma la data de entrada en datos de salida.
    IDataView dataTest = model.Transform(dataTestSet);
    //Inicia el proceso de evaluacion basandose en el algoritmo de clasificacion escogido
    CalibratedBinaryClassificationMetrics measurement = mlContext.BinaryClassification.Evaluate(dataTest, "Label");

    Console.WriteLine(".....");
    Console.WriteLine($"Precision: {measurement.Accuracy:P2}");
    Console.WriteLine($"AreaBajoLaCurva: {measurement.AreaUnderRocCurve:P2}");
    Console.WriteLine($"PuntajeF1: {measurement.F1Score:P2}");
    Console.WriteLine("-----Fin de evaluacion del modelo-----");
}
```

Ilustración 75. Evaluate [Propio]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Después de tener el código correspondiente al ingreso de la data, creación y entrenamiento del modelo de ML, y el método que evalúa a través de métricas especiales el correcto funcionamiento del modelo. Ahora se va a crear un pequeño método llamado **DataTest**, que recibe dos parámetros, el contexto ML y el modelo previamente entrenado. Allí se ingresa un texto aleatorio correspondiente a un error que se presenta en las pruebas automatizadas, y obtener la predicción de que tipo de error.
 - `PredictionEngine----` es una funcionalidad de ML.NET donde permite crear el ambiente para realizar las predicciones sencillas basándose en un modelo previamente entrenado, que hace uso de las clases llamadas **MistakeData**, y **MistakePrediction**, configuradas al comienzo de la codificación.
 - `Prediction---` variable donde se almacena la predicción que realice el algoritmo
 - `Model---` permite realizar operaciones haciendo uso del modelo entrenado
 - `CreatePredictionEngine---` funcionalidad que permite la creación de predicción para el texto ingresado haciendo uso del modelo entrenado, para ello se le deben pasar los parámetros de **MistakeData** y **MistakePrediction**, a su vez del modelo de entrenamiento.
 - Como se hace el ingreso de un texto aleatorio y totalmente nuevo, se realiza la creación de la variable `text`, la cual va a ser considerada como el nuevo valor de `MistakeText` en la clase **MistakeData**.
 - `resultPrediction----` variable que almacena el resultado de la predicción realizada por el modelo respecto al texto ingresado, es por ello que se hace el llamado del motor de predicción, la función `Predict` para realizar dicha acción y el texto que se ingresa al algoritmo.
 - Luego se muestra a través de la consola los resultados de la predicción y que tan probable es, los errores de sistema siempre manejan una probabilidad mayor a 0.5 mientras que las dadas por el programador son de baja probabilidad, para eso se hace uso de `if` para reafirmar el tipo de error que se da.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
//INGRESO DE DATA NUEVA PARA OBTENER PREDICION
1 referencia
private static void DataTest(MLContext mlContext, ITransformer model)
{
    //Hace uso del motor de prediccion dado, usando los datos, la funcion de prediccion y el modelo
    PredictionEngine<MistakeData, MistakePrediction> prediction = mlContext.Model.CreatePredictionEngine<MistakeData, MistakePrediction>(model);
    //Texto nuevo de entrada
    MistakeData text = new MistakeData
    {
        //MistakeText = "The HTTP request to the remote WebDriver server for URL http://localhost:7904/session/fe149868ee0f4bfe077d353406a25805/ele
        MistakeText= "El caso de prueba presento los siguientes errores: MSG El Nombre Emergente es incorrecto el esperado es Identificación Jefe y
    };
    //Resultado de la prediccion
    var resultPrediction = prediction.Predict(text);
    Console.WriteLine();
    Console.WriteLine("-----Prediccion-----");
    Console.WriteLine();
    Console.WriteLine($"Error: {resultPrediction.MistakeText} | Probabilidad: {resultPrediction.Probability} ");
    if (resultPrediction.Probability > 0.5)
    {
        Console.WriteLine($"Prediccion: {(Convert.ToBoolean(resultPrediction.Prediction) ? "" : "Error de sistema")}");
    }
    else
    {
        Console.WriteLine($"Prediccion: {(Convert.ToBoolean(resultPrediction.Prediction) ? "" : "Error causado por programador")}");
    }
    Console.WriteLine();
}
```

Ilustración 76. Predicción [Propia]

- Con todos los métodos necesarios creados, se procede a realizar el llamado de cada una de ellas en el orden adecuado en la función principal llamada **Main**.
 - Después de la creación del ML. context, se hace el llamado del método **LoadData** el cual maneja el escenario de TrainTestData ya que se fracciona un porcentaje de los datos totales, y se le da el parámetro que necesita para funcionar que es el contexto y se guarda en la variable `data` el valor que retorna LoadData.
 - Luego de tener la data necesaria, se hace el llamado del método **Model**, para realizar el respectiva construcción y entrenamiento del modelo, se le dan los dos parámetros que necesita para su funcionamiento tales como el contexto y la fracción de los datos totales destinados al entrenamiento a través de `data`. `TrainSet`, devolviendo el modelo listo el cual guarda en la variable llamada `model`.
 - El siguiente paso es evaluar el modelo haciendo el llamado del método **Evaluate**, dándole los parámetros necesarios tales como el contexto, el modelo entrenado, y la fracción de la data destinados para evaluar a través de `data`. `TestSet`. Mostrando las métricas de clasificación binaria que dan el porcentaje de precisión, verdaderos positivos y que tan preciso es el modelo.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Y finalmente se hace el llamado del método **DataTest**, el cual realiza la predicción del tipo de error basándose en un texto aleatorio, es por ello que para su funcionamiento se le dan dos parámetros, el contexto y el modelo entrenado y evaluado. Mostrando la predicción de que tipo de error se presentó.

```
static void Main(string[] args)
{
    // crea el contexto necesario para la configuracion del algoritmo.
    MLContext mlContext = new MLContext();
    //carga datos para entrenar en el contexto
    TrainTestData data = LoadData(mlContext);
    //Llama al modelo construido y entrenado
    ITransformer model = Model(mlContext, data.TrainSet);
    //Llama la funcion para evaluar el modelo creado
    Evaluate(mlContext, model, data.TestSet);
    //Llama la funcion para ingresar dato nuevo y obtener prediccion
    DataTest(mlContext, model);
}
```

Ilustración 77. Main [Propia]

FASE 4

4.1 Validar que la clasificación de los errores se realice adecuadamente.

En esta fase se hace uso de nuevas pruebas a las cuales se le realizaron actualizaciones que permitan dar resultados diferentes a los recibidos en las primeras pruebas lanzadas, para que así los errores obtenidos sean diferentes a los cuales se entrenaron y evaluar su funcionamiento de manera correcta.

Es por ello, que se realizó un lanzamiento masivo de pruebas nuevamente y se extrajeron los siguientes errores, tales como las ilustraciones 78,80,82 y 84 y se incluyeron en el sistema de clasificación obteniendo las predicciones para cada uno de ellos en el orden dado por las ilustraciones 79,81,83 y 85.

Error numero 1, presentado en la pruebas de automatización del programa Beneficios organizacionales en la sección de rechazo



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
✖ BeneficiosOrganizacionalesRechazo

Message: Error de Assert.Fail. 1546 :::::OpenQA.Selenium.WebDriverException: The
HTTP request to the remote WebDriver server for URL http://localhost:31134/
session/e71da7e87138f3bf2953e3d688f3ab5e/element timed out after 240 seconds.
--> System.Net.WebException: Anulada la solicitud: Se excedió el tiempo de espera
de la operación.
    en System.Net.HttpWebRequest.GetResponse()
    en OpenQA.Selenium.Remote.HttpCommandExecutor.MakeHttpRequest
(HttpRequestInfo requestInfo)
    --- Fin del seguimiento de la pila de la excepción interna ---
    en OpenQA.Selenium.Remote.HttpCommandExecutor.MakeHttpRequest
(HttpRequestInfo requestInfo)
    en OpenQA.Selenium.Remote.HttpCommandExecutor.Execute(Command
commandToExecute)
    en OpenQA.Selenium.Remote.DriverServiceCommandExecutor.Execute(Command
```

Ilustración 78. Error validación #1 [Propia]

Al ingresar el anterior error al sistema de clasificación se obtuvo una predicción, totalmente acertada, ya que ese error corresponde a un error de sistema, como se muestra en la siguiente imagen:

```
-----Prediccion-----
Error: : The HTTP request to the remote WebDriver server for URL http://localhost:31
134/session/e71da7e87138f3bf2953e3d688f3ab5e/element timed out after 240 seconds. --
-> System.Net.WebException: Anulada la solicitud: Se excedió el tiempo de espera de
la operación. | Probabilidad: 0,9981023
Prediccion: Error de sistema
```

Ilustración 79. Predicción algoritmo error #1 [Propia]

Error #2, correspondiente a la prueba automatizada del programa de préstamos sin interés, conocido internamente como BpSopreSi.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
✖ EnvioCorreoBpSopreSi  
  
Message: Error de Assert.Fail.  
1386 :::::OpenQA.Selenium.StaleElementReferenceException: stale element  
reference: element is not attached to the page document  
(Session info: chrome=96.0.4664.45)  
  en OpenQA.Selenium.Remote.RemoteWebDriver.UnpackAndThrowOnError  
(Response errorResponse)  
  en OpenQA.Selenium.Remote.RemoteWebDriver.Execute(String  
driverCommandToExecute, Dictionary`2 parameters)  
  en OpenQA.Selenium.Remote.RemoteWebElement.Execute(String  
commandToExecute, Dictionary`2 parameters)  
  en OpenQA.Selenium.Remote.RemoteWebElement.SendKeys(String text)  
  en APITest.APISelenium.SendKeys(String control, String value) en C:\Users\deisy\  
\Source\Repos\WEB_Kactus_Tests2\SeleniumAPITest\APISelenium.cs:lnea 141  
  en Web_Kactus_Test.SelfService_Nuevo1.EnvioCorreoBpSopreSi() en C:\Users  
\deisy\Source\Repos\WEB_Kactus_Tests2\Web_Kactus_Test
```

Ilustración 80. Error validación #2 [Propia]

La predicción dada por el sistema creado, correspondiente al error #2, se muestra a continuación, la cual es totalmente correcta al dar el error como error de sistema.

```
-----Prediccion-----  
Error: stale element reference: element is not attached to the page document (Session  
info: chrome = 96.0.4664.45) | Probabilidad: 0,9988234  
Prediccion: Error de sistema
```

Ilustración 81. Predicción error #2 [Propia]

Error #3 resultante de la prueba automatizada correspondiente a aprobación de permisos encargada por el jefe nivel específico.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
✖ PermisoSubFlujoTresAproAutoCargoNivelEspecifico  
  
Message: Error de Assert.Fail. 1774 ::::OpenQA.Selenium.NoSuchElementException:  
no such element: Unable to locate element: {"method":"xpath","selector":"//input  
[@id='ctl00_ContenidoPagina_txtMotivoCanc']"  
(Session info: chrome=96.0.4664.45)  
en OpenQA.Selenium.Remote.RemoteWebDriver.UnpackAndThrowOnError  
(Response errorResponse)  
en OpenQA.Selenium.Remote.RemoteWebDriver.Execute(String  
driverCommandToExecute, Dictionary`2 parameters)  
en OpenQA.Selenium.Remote.RemoteWebDriver.FindElement(String mechanism,  
String value)  
en OpenQA.Selenium.Remote.RemoteWebDriver.FindElementByXPath(String  
xpath)  
en OpenQA.Selenium.By.<>c__DisplayClass19_0.<XPath>b_0(ISearchContext  
context)  
en OpenQA.Selenium.By.FindElement(ISearchContext context)
```

Ilustración 82. Error validación #3 [Propia]

Como tercera validación del funcionamiento del sistema desarrollado, se obtuvo una respuesta correcta ya que ciertamente corresponde al mostrado a continuación:

```
-----Prediccion-----  
Error: no such element: Unable to locate element: {method:xpath:selector://input  
[@id='ctl00_ContenidoPagina_txtMotivoCanc']} | Probabilidad: 0,997654  
Prediccion: Error de sistema
```

Ilustración 83. Predicción error #3 [Propia]

Error #4 corresponde a seccion de pruebas conocida como NTC donde se examino el estado de la pagina web, conocido como frmLISolvía.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
SmartPeople_frmLISolvianTC
Message: Error de Assert.Fail.
2039 :Microsoft.VisualStudio.TestTools.UnitTesting.AssertFailedException: Error
de Assert.Fail. El caso de prueba presento los siguientes errores:
:MSG: El Nombre Emergente es incorrecto, el esperado es: Tipo de
Viatico y el encontrado es: Tipo de viatico
en Microsoft.VisualStudio.TestTools.UnitTesting.Assert.HandleFail(String
assertionName, String message, Object[] parameters)
en Microsoft.VisualStudio.TestTools.UnitTesting.Assert.Fail(String message)
en Web_Kactus_Test.SmartPeople_NTC_1.SmartPeople_frmLISolvianTC() en C:
\Users\deisy\Source\Repos\WEB_Kactus_Tests2\Web_Kactus_Test
\SmartPeople_NTC_1.cs:línea 17053
```

Ilustración 84. Error validación #4 [Propia]

Se obtiene la siguiente predicción, la cual concuerda rotundamente con el tipo de error presenciado, el cual corresponde al error generado por mi como programadora, siendo correcta la clasificación.

```
-----Prediccion-----
Error: . El caso de prueba presento los siguientes errores:
:MSG: El Nombre Emergente es incorrecto, el esperado es: Tipo de Viatico
y el encontrado es: Tipo de viatico | Probabilidad: 0,0008582077
Prediccion: Error causado por programador
```

Ilustración 85. Predicción error #4 [Propia]

Error #5 que corresponde también a la sección de NTC, específicamente del programa frmLINmJedes.



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
SmartPeople_frmLINmJedesNTC
Message: Error de Assert.Fail.
2009 :::::Microsoft.VisualStudio.TestTools.UnitTesting.AssertFailedException: Error
de Assert.Fail. El caso de prueba presento los siguientes errores:
::::::::::MSG: El Nombre Emergente es incorrecto, el esperado es: * Fecha
Comunicación y el encontrado es: * Fecha comunicación
en Microsoft.VisualStudio.TestTools.UnitTesting.Assert.HandleFail(String
assertionName, String message, Object[] parameters)
en Microsoft.VisualStudio.TestTools.UnitTesting.Assert.Fail(String message)
en Web_Kactus_Test.SmartPeople_NTC_1.SmartPeople_frmLINmJedesNTC() en C:
\Users\deisy\Source\Repos\WEB_Kactus_Tests2\Web_Kactus_Test
\SmartPeople_NTC_1.cs:línea 13033
```

Ilustración 86. Error validación #5 [Propia]

En esta última predicción ingresada al sistema se obtuvo una correcta predicción como se muestra a continuación, totalmente correcta, dando un balance de 5/5 clasificaciones de errores correctas por el sistema desarrollado.

```
-----Prediccion-----
Error: El caso de prueba presento los siguientes errores: ::::::::::::::::::::MSG: El Nombre
Emergente es incorrecto, el esperado es: *Fecha Comunicación y el encontrado es: *Fecha comunic
ación | Probabilidad: 0,0002892822
Predicción: Error causado por programador
```

Ilustración 87. Predicción error #5 [Propia]

Teniendo en cuenta lo anterior, se puede observar la gran precisión que tiene en algoritmo al momento de dar la respuesta a que tipo de error representa el ingresado, es por ello, que a través del método previamente diseñado llamado **Evaluate**, permite obtener tres mediciones importantes que evalúan el rendimiento del modelo en números, específicamente en porcentajes, predicción, area bajo la curva y una medición especial correspondiente a clasificación binaria como lo es F1, obteniendo lo siguiente.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
-----Evaluacion del modelo-----  
.....  
Precision: 98,98 %  
AreaBajoLaCurva: 100,00 %  
PuntajeF1: 0,98989898989899  
-----Fin de evaluacion del modelo-----
```

Ilustración 88. Evaluación del Modelo [Propia]

Para la evaluación del modelo, se hizo uso de tres métricas importantes al analizar sistemas de clasificación binaria, como lo son la precisión, el puntaje F1 y AUC. Estas métricas, como se expuso en el marco teórico, son basadas principalmente en el comportamiento de los verdaderos positivos o negativos vs los falsos positivos o negativos, permitiendo dar un balance en que tan preciso es el modelo, que tan capaz es de distinguir entre las dos etiquetas presentes si tener entre las predicciones falsos positivos o negativos.

- El rango de precisión se da entre 0 y 100%, donde el modelo puede ser considerado si maneja un rango mayor a 70% un modelo apto para su utilización, en este caso el modelo cumple con esas expectativas obteniendo un porcentaje de precisión del 98%, lo cual se noto al realizar las validaciones previas con errores aleatorios.
- El área bajo la curva es una medición que permite obtener la tasas de verdaderos positivos vs falsos positivos, es decir, mide que tan seguro es el algoritmo al momento de predecir evitando la presencia de predicciones erróneas, para que el modelo se considere aceptable y tenga valor, debe ser mayor al 50%. En este caso, al presenciar errores con etiquetas muy contundentes que permiten la clasificación clara y consisa de que error es, esto permitio al modelo tener alta seguridad dando un 100% en la toma de decisiones sin presencia de falsos positivos.
- Puntaje F1, da el valor de precisión que tiene el modelo entre un valor de 0 a 1, entre mas cercano a 1, su precisión es mas alta, en sistemas de clasificación binaria, complementando el rango de precisión en porcentaje de la primera medición, en este caso su valor fue de 0.98 aproximadamente, dando credibilidad de lo mostrado en la primera métrica y las validaciones realizadas.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



CAPITULO VI. RESULTADOS

Ante las necesidades del equipo de automatización en la compañía Digital Ware al realizar el proceso de ejecución, revision y soporte de pruebas automatizadas teniendo en cuenta la gran cantidad que se ejecutan diariamente, semanalmente y mensualmente de manera masiva, se creó un sistema de clasificación de errores basándose en Machine Learning que permitió las predicciones de los errores presentados en las pruebas automatizadas, específicamente del proyecto web que maneje y controle, donde se obtuvo una respuesta positiva con resultados altamente precisos dando validez al uso de este sistema dentro del equipo permitiendo así facilitar el reajuste de pruebas haciendo uso del menor tiempo posible.

Adquirir todos los conocimientos necesarios para la aplicación de machine Learning en pruebas automatizadas tuvo una influencia muy importante en el correcto funcionamiento del sistema creado, donde la extracción de las clases de modelos presentes, cuales eran cada una de sus tecnicas y como era el comportamiento a nivel matemático de los algoritmos, hizo que la elección del aprendizaje supervisado, con la técnica de clasificacion binaria y el algoritmo de regresión logística fuera totalmente correcta y acertada que con el uso de herramientas .NET, este sistema puede ser acoplado al proyecto raíz de la empresa ya que se cuenta con toda la compatibilidad necesaria ya que se usó el mismo lenguaje, misma interfaz de trabajo, solo es necesarios NuGets que fácilmente pueden ser implementados en las soluciones sin afectar las configuraciones que ya se tengan, dando la base necesaria de confiabilidad en uso de librerías como ML.NET para proyectos de aprendizaje automatizado en el futuro de la empresa.

La implementación de la tecnología Azure permitió una mejor organización de los aspectos importantes que conforman las pruebas automatizadas, que al funcionar en conjunto, permiten obtener sus resultados de manera legible con uso de gráficas obteniendo con exactitud los porcentajes correspondientes las pruebas que obtuvieron errores, especificar cuál fue el error presentado, facilitando la extracción correcta de todos estos errores necesarios para el entrenamiento del algoritmo de manera sencilla y concreta, y como son compilaciones programadas, al finalizarse su ejecución alertan de presencia de más datos que pueden fueron usados para nuevos entrenamiento, evaluación o predicción, teniendo variedad de información dando una correcta evaluación de la precisión del sistema creado. En conjunto con la tecnología de ML.NET, al manejar formatos específicos para la organización de los datos usados para el entrenamiento, facilito la adecuación de la información extraída de los pipelines, a una estructura que al usar los escenarios que tiene ML.NET para manejo de data, tuviese esa facilidad de identificar cual era el error y que etiqueta tenía cada uno, este aspecto al ser desarrollado



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



correctamente implicó mucho en el alto porcentaje de precisión que obtuvo el sistema de clasificación ya que se fue claro y conciso de lo que se requería.

La creación del sistema al ser dividido por secciones a nivel de codificación, y en el orden adecuado, permitió llevar una estructura ordenada que facilitó el entendimiento del código a su vez analizar el paso a paso interno que realiza el código. Como primer aspecto se realizó la creación del Contexto, el cual al ejecutarse obtiene un resultado positivo, ya que, en él, se realiza toda la edificación del sistema, como se observa allí se almacenaron todas las instrucciones usadas en el código tales como la clasificación binaria, los modelos, las transformaciones, manejar de data, etc. Siendo la base más importante de ML.NET.

Nombre	Valor	Tipo
miContext	{Microsoft.ML.MLContext}	Microsoft.ML.MLConte...
▶ AnomalyDetection	{Microsoft.ML.AnomalyDetectionCatalog}	Microsoft.ML.AnomalyDet...
▶ BinaryClassification	{Microsoft.ML.BinaryClassificationCatalog}	Microsoft.ML.BinaryClassif...
▶ Clustering	{Microsoft.ML.ClusteringCatalog}	Microsoft.ML.ClusteringCa...
▶ ComponentCatalog	{Microsoft.ML.Runtime.ComponentCatalog}	Microsoft.ML.Runtime.Co...
▶ Data	{Microsoft.ML.DataOperationsCatalog}	Microsoft.ML.DataOperati...
▶ Forecasting	{Microsoft.ML.ForecastingCatalog}	Microsoft.ML.ForecastingC...
▶ Model	{Microsoft.ML.ModelOperationsCatalog}	Microsoft.ML.ModelOpera...
▶ MulticlassClassification	{Microsoft.ML.MulticlassClassificationCatalog}	Microsoft.ML.MulticlassCl...
▶ Ranking	{Microsoft.ML.RankingCatalog}	Microsoft.ML.RankingCata...
▶ Regression	{Microsoft.ML.RegistrationCatalog}	Microsoft.ML.RegistrationC...
▶ TempFilePath	"C:\\Users\\deka1\\AppData\\Local\\Temp\\"	string
▶ Transforms	{Microsoft.ML.TransformsCatalog}	Microsoft.ML.TransformsC...
▶ Non-Public members		

Ilustración 89. Resultado Contexto [Propia]

Seguido de ello se muestra cómo se fracciona la data total en dos partes, una para entrenamiento y otra para evaluación, para la primera opción, se da el resultado de lo indicado a través del código, donde identifica las columnas que tiene, cual es el tipo de variable que almacena, en donde esta almacenada la etiqueta, que posición de la columna, que se ingresa y que se retorna.

- Muestra de que ruta se hizo el llamado de la data acoplada con el formato de ML.NET, validando que adquiere la información del .txt correcto.

Nombre	Valor	Tipo
dataView	null	Microsoft.ML.IDat...
load	"C:\\Users\\deka1\\source\\repos\\TesisErroresAutomatizacion\\TesisErroresAutomatizaci..."	string
miContext.Data	{Microsoft.ML.DataOperationsCatalog}	Microsoft.ML.Data...

Ilustración 90. Ruta data [Propia]



SC-CER96940



"Formando líderes para la construcción de un nuevo país en paz"

Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
 www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- Valida que la ubicación de la etiqueta en el archivo .txt fue la correcta, para ello muestra cómo se conforman los datos cargados, en que posición están a nivel de columnas y que tipo de variable almacena cada una de ellas, las cuales coinciden con lo codificado.

dataView	{Microsoft.ML.Data.TextLoader.BoundLoader}	Microsoft.ML.IDat...
CanShuffle	false	bool
Schema	{2 columns}	Microsoft.ML.Data..
[0]	{MistakeText: String}	Microsoft.ML.Data..
[1]	{Label: Boolean}	Microsoft.ML.Data..
Raw View		

Ilustración 91. Data composición [Propia]

- Realiza la correcta fracción de los datos, los cuales son divididos en uno para entrenamiento y otro para evaluación del algoritmo.

data	{Microsoft.ML.DataOperationsCatalog.TrainTestData}	Microsoft.ML.Data...
TestSet	{Microsoft.ML.Transforms.ColumnSelectingTransformer.SelectColumnsDataTransform}	Microsoft.ML.IDat...
TrainSet	{Microsoft.ML.Transforms.ColumnSelectingTransformer.SelectColumnsDataTransform}	Microsoft.ML.IDat...

Ilustración 92. Fracción Data [Propia]

Validando que la información que recibe el sistema, con la información clara y concisa, a través de la consola se indica en qué momento inicia y en qué momento termina, tanto la carga de los datos como el entrenamiento del modelo, obteniendo una respuesta muy rápida que se reduce en este caso a segundos, confirmando que el uso de la clasificación binaria en la técnica de regresión logística se comportó adecuadamente al no presentar errores en la construcción del modelo dándole 100% funcional.

```
Consola de depuración de Microsoft Visual Studio
-----Inicio Carga de data-----
-----Carga completa de data-----
-----Inicio de construccion del modelo-----
-----Finalizacion de construccion del modelo-----
```

Ilustración 93. Validación carga y entreno correcto del modelo [Propia]

El tiempo de ejecución varía en torno a la cantidad de datos que se manejarán y la técnica aplicada, de ahí vale la correcta elección del algoritmo que permitieron que el sistema construyera este modelo sin exceder mucho en tiempo y a su vez ahorrar recursos evitando saturaciones en la ejecución del mismo. Como se observa en las siguientes gráficas, los recursos que consume



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



a nivel de memoria y CPU son muy bajos, siendo de memoria un máximo de 40MB y un pico máximo de consumo de CPU menor al 30% del total de todos los procesadores.

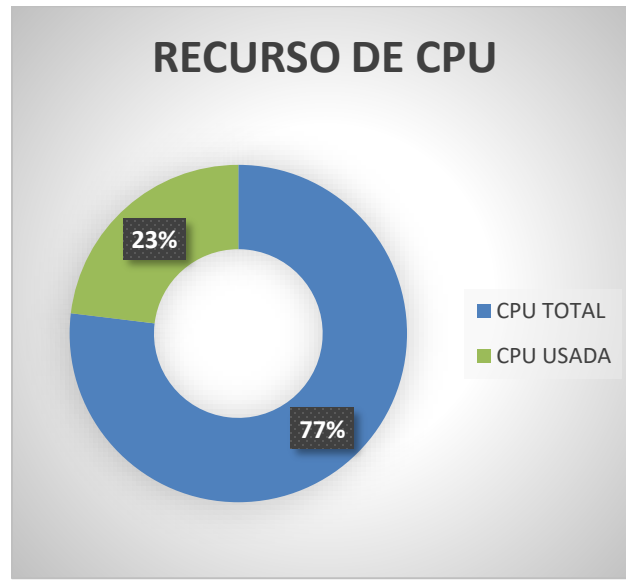


Ilustración 94. Consumo CPU [Propia]

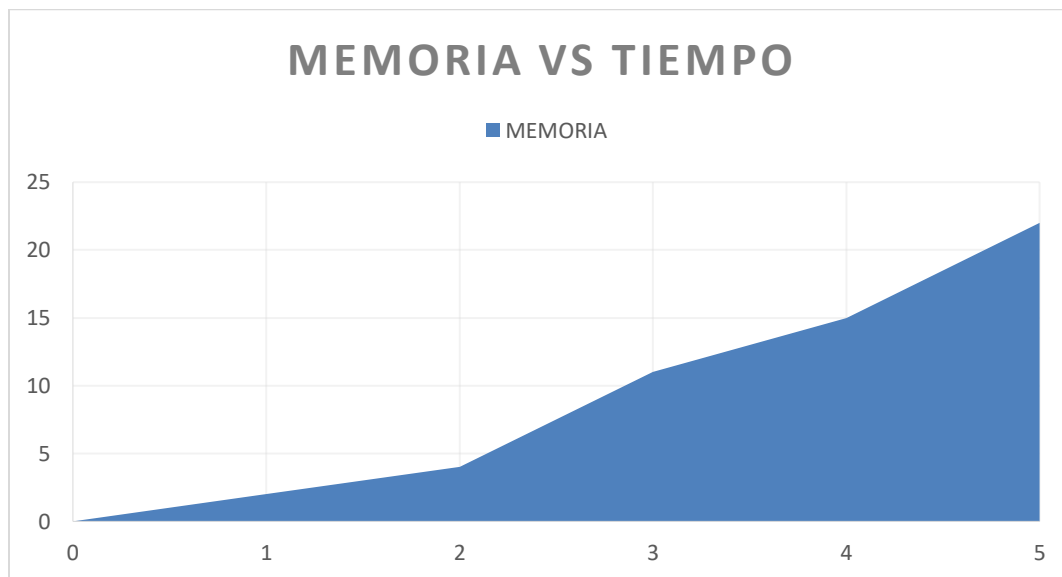


Ilustración 95. Memoria vs Tiempo [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

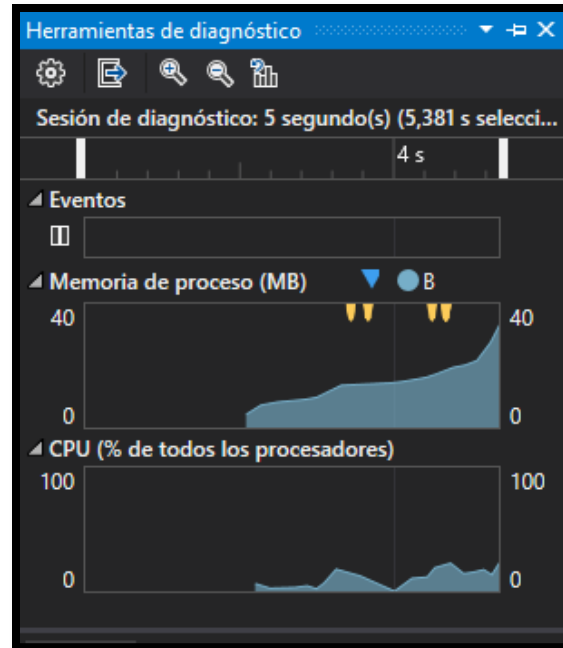


Ilustración 96. Consumo de recursos

Siendo así un sistema muy trabajable a nivel de gastos de recursos, al alcance de todos y con una tecnología como ML.NET que tiene un alto grado de precisión y siendo una de las mejores tecnologías para machine Learning en la actualidad y futuro para innovación de nuevos sistemas de aprendizaje automático. Al ejecutar el sistema de manera completa se pueden observar cada una de las etapas indicando a través de la consola, el proceso que se está llevando en paralelo a la ejecución del código, por lo cual se observa cuando ya se realizó la carga de los datos, cuando se inició y finalizó la construcción del modelo, obtener las métricas del modelo gracias a la evaluación que se le hace al modelo haciendo uso de datos diferentes a los usados para el entrenamiento, y finalmente se obtiene la predicción del error ingresado para realizar la clasificación. En la imagen se hizo el ingreso de un error presentado en nuevas pruebas automatizadas al sistema dando la predicción correcta al tipo de error, ya que correspondía a un error que como programadores generamos para la calidad del producto de Digital Ware.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



```
Consola de depuración de Microsoft Visual Studio
-----Inicio Carga de data-----
-----Carga completa de data-----
-----Inicio de construccion del modelo-----
-----Finalizacion de construccion del modelo-----
-----Evaluacion del modelo-----
.....
Precision: 98,98 %
AreaBajoLaCurva: 100,00 %
PuntajeF1: 0,98989898989899
-----Fin de evaluacion del modelo-----

-----Prediccion-----

Error: . El caso de prueba presento los siguientes errores::::::::::::::::::MSG: El Nombre Emergente es incorrecto,
el esperado es: Tipo de Viatico y el encontrado es: Tipo de viatico | Probabilidad: 0,00087751675
Prediccion: Error causado por programador
```

Ilustración 97. Visualización de la predicción [Propia]

Al ser un sistema de clasificación con un tiempo bajo de respuesta y procesamiento con bajos recursos, aplica exactamente al propósito del desarrollo del mismo, hacer que el equipo de automatización pueda disminuir el tiempo usado en reajuste y corrección de pruebas, agilizando el proceso evitando saturaciones en el correcto flujo de la empresa ante la evaluación de calidad.

A nivel de validaciones, las métricas destinadas al modelo de clasificación creado, determina cuan satisfactorio es el uso del sistema de clasificación creado, ya que, a través de estas, se analizó la efectividad con la que responde el sistema ante las predicciones que realizo, en el caso actual, dando como resultado un buen balance entre precisión y calidad de predicciones, se obtuvo lo siguiente.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

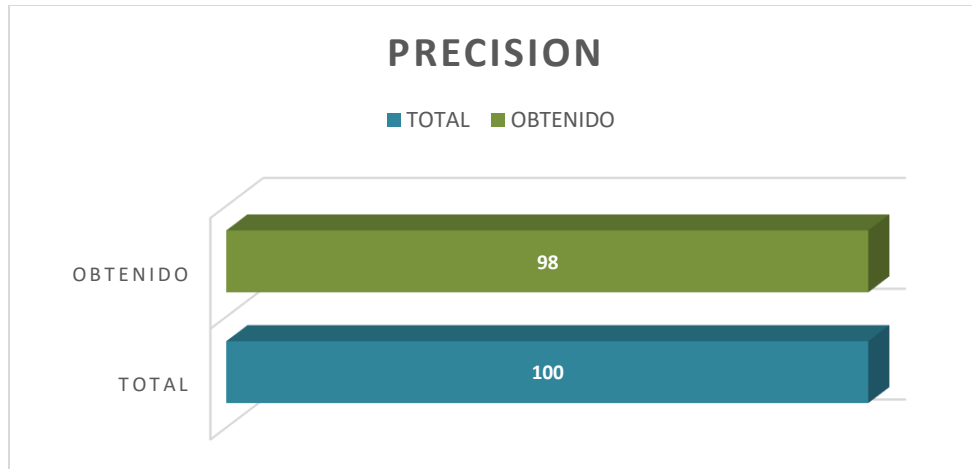


Ilustración 98. Precisión sistema [Propia]

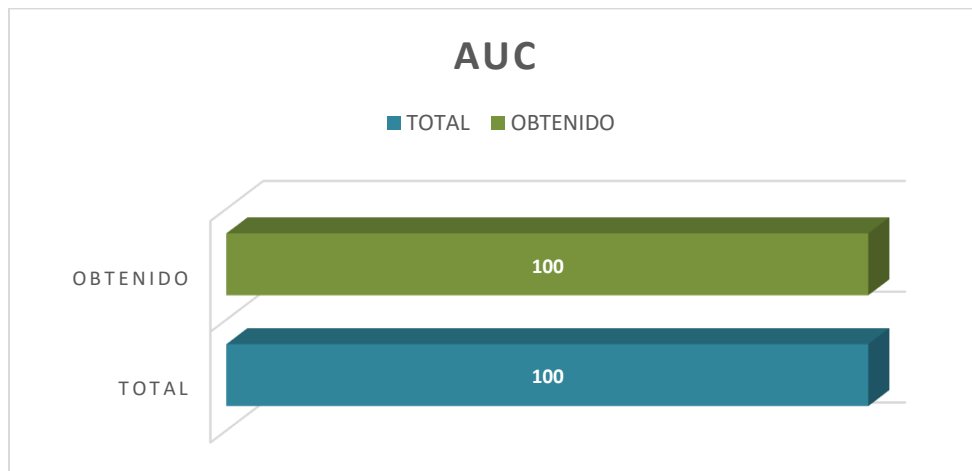


Ilustración 99. Area bajo la curva [Propia]



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!

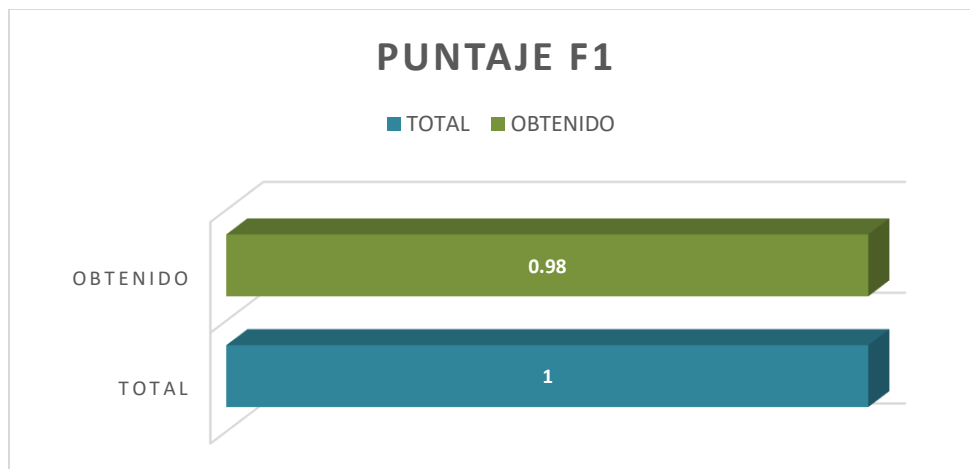


Ilustración 100. Puntaje F1 [Propia]

Tabla 3. Resultados en métricas [Propia]

PRECISION	CONTUNDENCIA	RANGO DE PREDICCIÓN
Rango de 50 a 100% para considerar el modelo apto, se obtuvo un porcentaje del 98% de precisión.	Rango de 0 a 100% donde se mide la claridad de las predicciones evitando la presencia de falsos positivos fue de un 100%	Escala de 0 a 1, que tan preciso estuvo el modelo creado se obtuvo como resultado un valor de 0.98 , confirmando el porcentaje obtenido a nivel de precisión.

Estos resultados positivos obtenidos, se basan en la correcta estructuración de la data que se utilizó para el entrenamiento, ya que el algoritmo identificó fácilmente la etiqueta y el error, por la organización que llevaba el formato de separación por tabulaciones que se configuró previamente y con el manejo de data en ML.NET se obtuvo la información correcta dando un balance realmente positivo en la creación de este sistema catalogándolo como un sistema altamente preciso y confiable.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
 Pamplona - Norte de Santander - Colombia
 Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE

¡Seguimos avanzando!



CAPITULO VII. CONCLUSIONES

El desarrollo de este sistema usando modelos y técnicas de Machine Learning, con uso de herramientas de una de las tecnologías más importantes a nivel de Microsoft como es .NET fue exitoso en todo sentido, en la elección de las herramientas dando total compatibilidad con las que se usan en la empresa en el área de automatización, tales como Visual, lenguaje C#, Azure, Selenium, la elección de las librerías de ML.NET para la aplicación de inteligencia artificial por su facilidad de entendimiento y llamado de las funciones globales que permitieron la codificación ordenada y estructurada, se hace uso de un modelo y una técnica correspondiente, como lo es clasificación binaria con la técnica de regresión logística, formando el equipo perfecto para el desarrollo final, en comparación con los demás modelos expuestos que se enfocaban en objetivos como aprendizajes al error, o aprendizaje sin etiquetas, no siendo el objetivo de este sistema.

Para el tratamiento de la data y toda la parametrización previa que requiere antes de ser usada para el entrenamiento del sistema, al elegir las librerías de ML.NET, da varios formatos guías que permitieron estructurar la información de tal manera que se viera ordenada y precisa para que al momento de cargarla se pudo distinguir de manera automática que filas y columnas conformaban el archivo de texto, gracias a la separación de tabulares y a las condiciones de código donde se daba la posición donde se encontraba la etiqueta o texto. Por lo cual este paso debe tener el mayor de los cuidados, ya que la manera en la que se distribuya la información va a determinar la precisión que tiene sistema de clasificación, por lo tanto, es indispensable que entre la cantidad de datos de una etiqueta 0 y datos de la etiqueta 1 sea la misma para evitar resultados inestables.

El uso de librerías ML.NET en la codificación y de NET.CORE en el espacio de trabajo, se complementaron de gran manera permitiendo la creación de un sistema muy completo a nivel de Machine Learning en herramientas de .NET con lenguajes de programación variados para facilitar al programador el proceso de creación teniendo la posibilidad de escoger el que más conozca e implementarlo, es este caso, por compatibilidad empresarial, se hizo uso de C# y el uso de aplicación de consola permitió el ahorro de muchos recursos de máquina y rápida respuesta del paso a paso que realizó el sistema.

Este proyecto se usará como base en la creación de varios sistemas de Machine Learning que sigan beneficiando la integración de aprendizaje en sistemas internos de la compañía, aclarando que se plantea ser compatible con las herramientas que se usan para la generación de pruebas automatizadas a nivel de Ophelia en Kactus, por lo cual este proyecto realizado facilitará la continuación del planteado por el equipo de automatización a largo plazo, debido a la fiabilidad



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



que genera el uso de tecnología .NET y el alto porcentaje de efectividad resultante, y su compatibilidad a largo plazo con el lenguaje C# y Visual Studio, que fueron usadas en su totalidad para la creación de este sistema de clasificación.

Al implementar el sistema de clasificación se realiza la evaluación del modelo con métricas estipuladas para clasificación binaria donde se consiguió las correctas predicciones acerca del tipo de error ingresado, teniendo un porcentaje de acierto mayor al 90 %, el cual toma como muestra la mitad de la data de entrada, la cual se secciono para evaluación del modelo, teniendo cerca de 100 errores para cumplir este propósito, obteniendo un 98% siendo un modelo de clasificación de alta efectividad, el cual es convalidado con la contundencia del 100% ante la no presencia de falsos positivos, siendo un sistema de clasificación confiable y efectivo. Este alto porcentaje de debe también gracias a la correcta elección del algoritmo de clasificación binaria y el tratamiento de la data de entrenamiento la cual debe tener una división de errores equilibrada donde se presenten errores de sistema de igual manera que se presenten errores generados por nosotros, además de la estructuración que se uso para la correcta lectura del sistema.

A nivel de consumo de tiempo y recursos de máquina, el sistema de clasificación creado tiene un punto positivo ante estos dos aspectos, ya que al ser configurado con tecnología .NET permite que el ahorro de recursos tanto de memoria como de CPU sea muy bajo, específicamente de máximo 50Mb en memoria y menos del 30% en CPU siendo relativamente bajo en comparación a las especificaciones técnicas que manejan las maquinas de la empresa, por lo cual no genera sobrecargas y se puede seguir usando la maquina sin problemas. A nivel de tiempo, usar este sistema de clasificación obtiene una respuesta rápida, ayudando a cumplir el objetivo de ahorrar tiempo dedicado a reajuste y corrección de errores de pruebas, teniendo una respuesta casi instantánea por parte del algoritmo, si se habla de tiempo de entrenamiento, depende de la cantidad de datos ingresados y de objetivo final del modelo, en este caso al ser una clasificación de solo dos estados y tener información muy bien estructurada, se tardo aproximadamente 5 segundos siendo además de efectivo muy fluido y rápido su procesamiento.



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



REFERENCIAS BIBLIOGRAFICAS

- [1] Digital Ware (2020). *Digital Ware, Tecnología que cambia Vida de las personas. SOBRE NOSOTROS.* <https://www.digitalware.com.co/sobre-nosotros/> (Consultado: 7 noviembre 2021).
- [2] Digital Ware (2020). *Intranet. INDUCCION CORPORATIVA.* http://intranet.digitalware.co/images/noticias/internas/Gestion_Humana/2021/Inducci%C3%B3n_corporativa_06_09_2021.pdf (Consultado: 7 noviembre 2021).
- [3] RIPISPA (2019). *Ripispa, Tecnología al servicio de la industria. AUTOMATIZACION.* <https://ripipsabots.com/automatizacion/> (Consultado: 7 noviembre 2021).
- [4] RED HAT (2015). *Red Hat Forum Madrid. Automatization de procesos. Que es la automatizacion.* <https://www.redhat.com/es/topics/automation/whats-it-automation> (Consultado: 7 noviembre 2021).
- [5] Gary Briceño (30 de marzo de 2020). *Club de Tecnología, Conocimiento es poder. Pruebas Automatizadas: tipos y conceptos erróneos.* <https://www.clubdetecnologia.net/blog/2020/pruebas-automatizadas-tipos-y-conceptos-erroneos/> (Consultado: 7 noviembre 2021).
- [6] Programación y Mas. *Blog. Buenas prácticas. Diferentes tipos de testing en el Desarrollo de software.* <https://programacionymas.com/blog/tipos-de-testing-en-desarrollo-de-software> (Consultado: 7 noviembre 2021).
- [7] YeePLY (2020). *YeePLY, tendencias y habilidades. Que son las pruebas unitarias y como llevar una a cabo.* <https://www.yeeply.com/blog/que-son-pruebas-unitarias/> (Consultado: 7 noviembre 2021).
- [8] Datcom-monitor (octubre 16, 2020). *Loadview. Tipos de pruebas de software: diferencias y ejemplos. h* <https://www.loadview-testing.com/es/blog/tipos-de-pruebas-de-software-diferencias-y-ejemplos/> (Consultado: 7 noviembre 2021).
- [9] Wikipedia (2018). *Wikipedia, la enciclopedia libre. Pruebas de software.* https://es.wikipedia.org/wiki/Pruebas_de_software#Pruebas_funcionales (Consultado: 7 noviembre 2021).
- [10] Sten Pittet (2016). *Atlassian CI/CD. Los distintos tipos de pruebas de software.* <https://www.atlassian.com/es/continuous-delivery/software-testing/types-of-software-testing> (Consultado: 7 noviembre 2021).
- [11] Brian (2018). *It-swarm-es.com. Diferencias entre pruebas funcional y pruebas Punta a Punta.* <https://www.it-swarm-es.com/es/testing/diferencia-entre-prueba-funcional-y-prueba-de-punta->



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



punta/835959775/ (Consultado: 7 noviembre 2021).

[12] Manuel Cillero (2017). *Mi circunstancia digital. Pruebas de Regresión.* <https://manuel.cillero.es/doc/metodologia/metrica-3/tecnicas/pruebas/regresion/> (Consultado: 7 noviembre 2021).

[13] Vipin Joshi (2021). *Cynoteck. Que es la pruebas de humo. Su función y beneficios.* <https://cynoteck.com/es/blog-post/what-is-smoke-testing/> (Consultado: 7 noviembre 2021).

[14] Juan Pablo Monsalve (2020). *YouTube. Que es Selenium. Automatización de pruebas.* <https://www.youtube.com/watch?v=PtmOR7GaAHI> (Consultado: 7 noviembre 2021).

[15] Fernando Coelho (2019). *Digital 55. Herramientas de testing: introducción a Selenium.* <https://www.digital55.com/desarrollo-tecnologia/herramientas-testing-introduccion-selenium/> (Consultado: 7 noviembre 2021).

[16] Reyes (2021). *Programa en Línea. Selenium y la automatización de las pruebas.* <https://www.programaenlinea.net/selenium-y-la-automatizacion-de-las-pruebas/> (Consultado: 7 noviembre 2021).

[17] Netec. *Netec, Expertos enseñando a expertos. Que es Oracle.* <https://www.netec.com/que-es-oracle> (Consultado: 7 noviembre 2021).

[18] Tic Portal (2019). *Tic. Portal. Base de datos SQL.* <https://www.ticportal.es/glosario-tic/base-datos-sql> (Consultado: 7 noviembre 2021).

[19] Neuronet (2014). *Bases de datos. Licenciamiento. Licenciamiento Oracle.* <https://neuronet.cl/licenciamiento-oracle/> (Consultado: 7 noviembre 2021).

[20] Microsoft (2021). *Microsoft Guía de C#. Paseo por el lenguaje C#.* <https://docs.microsoft.com/es-es/dotnet/csharp/tour-of-csharp/> (Consultado: 7 noviembre 2021).

[21] BeSoftware. *Soluciones TIC BSW. Que es C# y para que sirve.* <https://bsw.es/que-es-c/> (Consultado: 7 noviembre 2021).

[22] JorTurFer (2019). *Fixed Buffer. Blog de divulgación sobre tecnologías de Microsoft. La potencia de la Reflexión en C#.* <https://www.fixedbuffer.com/la-potencia-de-la-reflexion-en-c-parte-2-ensamblados/> (Consultado: 7 noviembre 2021).

[23] Microsoft (2021). *Microsoft Net Framework. Introducción a .NET Framework.* <https://docs.microsoft.com/es-es/dotnet/framework/get-started/> (Consultado: 7 noviembre 2021).

[24] Antonio Torres (2008). *Slideshare. Tecnología Microsoft.Net.*



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



<https://es.slideshare.net/dixzan/tecnologia-microsoft-net> (Consultado: 7 noviembre 2021).

[25] Henry Bonilla Castro (2010). *Desarrollo en .NET. Arquitectura del Framework .NET.* <https://boninet.wordpress.com/2010/08/29/arquitectura-del-framework-net/> (Consultado: 7 noviembre 2021).

[26] Cesar Diaz Alcolea (2020). *Open Webinars. Que es .NET Core.* <https://openwebinars.net/blog/que-es-net-core/> (Consultado: 7 noviembre 2021).

[27] Cesar Velarde (2020). *Globant Stay Relevant. NET 5 Framework unificado.* <https://stayrelevant.globant.com/es/net5-framework-unificado/> (Consultado: 7 noviembre 2021).

[28] NetApp. *NetApp. Que es la inteligencia artificial.* <https://www.netapp.com/es/artificial-intelligence/what-is-artificial-intelligence/> (Consultado: 7 noviembre 2021).

[29] Sebastian Molinetti (2019). *Telefonica. Think Big Empresas. Conoce los tipos de inteligencia artificial empresariales.* <https://empresas.blogthinkbig.com/tipos-inteligencia-artificial-empresariales/> (Consultado: 7 noviembre 2021).

[30] Redacción APD (2021). *APD. Metodos y tecnicas de inteligencia artificial: Cuales son y para que se usan.* <https://www.apd.es/tecnicas-de-la-inteligencia-artificial-cuales-son-y-para-que-se-utilizan/> (Consultado: 8 noviembre 2021).

[31] Fundación Telefónica (2019). *Espacio. Las posibilidades actuales de la inteligencia artificial.* <https://espacio.fundaciontelefonica.com/noticia/las-posibilidades-de-la-inteligencia-artificial/> (Consultado: 8 noviembre 2021).

[32] Andrés Gonzales. *Clever data. Machine Learning.* <https://cleverdata.io/que-es-machine-learning-big-data/> (Consultado: 8 noviembre 2021).

[33] Redacción APD (2019). *APD. Que es Machine Learning y como funciona.* <https://www.apd.es/que-es-machine-learning/> (Consultado: 8 noviembre 2021).

[34] Datos Gob Es (2020). *Innovacion. Como aprenden las maquinas. Machine Learning y sus Diferentes tipos.* <https://datos.gob.es/es/blog/como-aprenden-las-maquinas-machine-learning-y-sus-diferentes-tipos> (Consultado: 8 noviembre 2021).

[35] Daniela Rodríguez Galindo (2021). *CII.IA. Aprendizaje Supervisado.* <https://www.ciiia.mx/noticiasciiia/aprendizaje-supervisado-1> (Consultado: 8 noviembre 2021).

[36] Rafael Zambrano (2019). *OpenWebinars. Modelos de Machine Learning.* <https://openwebinars.net/blog/modelos-de-machine-learning/> (Consultado: 8 noviembre 2021).



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



- [37] Ampatishan Sivalingam (2019). Start it up. What is machine Learning. <https://medium.com/swlh/what-is-machine-learning-9b569ff7858a> (Consultado: 8 noviembre 2021).
- [38] Content Writer GoTD (2019). The Learning Company got digital. Inteligencia Artificial_ un enfoque para no tecnólogos. <https://gotodigital.es/inteligencia-artificial-un-enfoque-para-no-tecnologos/> (Consultado: 8 noviembre 2021).
- [39] José Heras (2020). IArtificial. Regresión Logística para clasificación. <https://www.iartificial.net/regresion-logistica-para-clasificacion/> (Consultado: 8 noviembre 2021).
- [40] José Heras (2020). IArtificial. Árboles de decisión con ejemplos. https://www.iartificial.net/arboles-de-decision-con-ejemplos-en-python/#Arboles_de_Decision_para_Clasificacion (Consultado: 8 noviembre 2021).
- [41] José Heras (2020). IArtificial. Máquinas de vectores de soporte. <https://www.iartificial.net/maquinas-de-vectores-de-soporte-svm/> (Consultado: 8 noviembre 2021).
- [42] José Heras (2020). IArtificial. Clasificación o Regresión. <https://www.iartificial.net/clasificacion-o-regresion/> (Consultado: 8 noviembre 2021).
- [43] José Heras (2020). IArtificial. Regresión Lineal: teoría y ejemplos. <https://www.iartificial.net/regresion-lineal-con-ejemplos-en-python/> (Consultado: 8 noviembre 2021).
- [44] Wikipedia (2021). Wikipedia, la enciclopedia libre. Regresión no Lineal. https://es.wikipedia.org/wiki/Regresi%C3%B3n_no_lineal (Consultado: 8 noviembre 2021).
- [45] Diego Calvo (2019). Diego Calvo. Aprendizaje no supervisado. <https://www.diegocalvo.es/aprendizaje-no-supervisado/> (Consultado: 8 noviembre 2021).
- [46] Michael Gogging (2018). Enstoa. Machine Learning in Construction: How Clustering Data Can Improve Processes (part 2 of 2). <https://enstoa.com/blog/machine-learning-construction-how-clustering-data-can-improve-processes-part-2-of-2> (Consultado: 8 noviembre 2021).
- [47] Jordi Olle (2020). Conceptos claros. Qué es y para qué sirve el clustering – un ejemplo de aplicación práctico. <https://conceptosclaros.com/que-es-clustering/> (Consultado: 8 noviembre 2021).
- [48] Ainhoa Lafuente (2018). Aukera mastering data. Reducción de la dimensionalidad. <https://aukera.es/blog/reduccion-dimensionalidad/> (Consultado: 8 noviembre 2021).
- [49] Ligdi González (2020). Aprende IA. Reducción de la dimensionalidad. <https://aprendeia.com/reduccion-de-la-dimensionalidad-machine-learning/> (Consultado: 8 noviembre 2021).



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



[50] Na8 (2020). *Aprende Machine Learning en Español. Aprendizaje por refuerzo.* <https://www.aprendemachinelearning.com/aprendizaje-por-refuerzo/> (Consultado: 8 noviembre 2021).

[51] Linux Post Install. *Desde Linux. .NET y ML.NET: Plataformas de Código Abierto de Microsoft.* <https://blog.desdelinux.net/net-ml-net-plataformas-codigo-abierto-microsoft/> (Consultado: 9 noviembre 2021).

[52] Mayank (2019). *CIS Make it possible. Una breve introducción técnica de ML.NET.* <https://www.cisin.com/coffee-break/es/enterprise/a-brief-technical-introduction-of-ml-net.html> (Consultado: 9 noviembre 2021).

[53] Microsoft (2021). *Microsoft Machine Learning. ML.NET.* <https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet> (Consultado: 10 noviembre 2021).

[54] Microsoft (2021). *Microsoft NET. What is ML.NET and how does it work.* <https://docs.microsoft.com/en-us/dotnet/machine-learning/how-does-ml-dotnet-work> (Consultado: 10 noviembre 2021).

[55] Fabian Farola, Carlos Iván Chesñevar, Carlos Alvez, Graciela Etchart, Ernesto, Miranda Silvia Ruiz, Juan José Aguirre, Juan Carlos Teze. (2017). *Universidad Nacional del Sur, Argentina. FRAMEWORK SDF MACHINE LEARNING EN TRANSACCIONES FINANCIERAS Y DETECCIÓN TEMPRANA DE FRAUDES.* (Consultado: 9 noviembre 2021).

[56] Edwar Fabian León Muñoz (2020). *Universidad Nacional de Santa, Perú. MODELO DE MACHINE LEARNING PARA LA CLASIFICACIÓN DE ESTUDIANTES DE ACUERDO A SU RENDIMIENTO ACADÉMICO EN EL CENTRO DE IDIOMAS DE LA UNIVERSIDAD NACIONAL DEL SANTA.* (Consultado: 9 noviembre 2021).

[57] Arles Felipe García Maya (2020). *Estudio comparativo de técnicas de clasificación binaria con múltiples anotadores.* <https://repositorio.unal.edu.co/bitstream/handle/unal/75807/1088294092.2019.pdf?sequence=4&isAllowed=y> (Consultado: 9 noviembre 2021).

[58] Amazon Web Services. *Clasificación Binaria.* https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/binary-classification.html (Consultado: 8 diciembre 2021).

[59] Víctor Román (2019). *Ciencia y datos. Aprendizaje Supervisado: Introducción a la Clasificación y Principales Algoritmos.* <https://medium.com/datos-y-ciencia/aprendizaje-supervisado-introducci%C3%B3n-a-la-clasificaci%C3%B3n-y-principales-algoritmos-dadee99c9407> (Consultado: 9 diciembre 2021).

[60] José Martínez Heras (2020). *IArtificial.net. precisión, Recall, F1, Accuracy en clasificación.*



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”

Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co



ACREDITADA INSTITUCIONALMENTE
¡Seguimos avanzando!



<https://www.iartificial.net/precision-recall-f1-accuracy-en-clasificacion/> (Consultado: 9 diciembre 2021).

[61] Víctor Román (2019). *Ciencia y datos. Machine Learning Supervisado: Fundamentos de la Regresión Lineal*. <https://medium.com/datos-y-ciencia/machine-learning-supervisado-fundamentos-de-la-regresi%C3%B3n-lineal-bbcb07fe7fd> (Consultado: 9 diciembre 2021).



SC-CER96940



“Formando líderes para la construcción de un nuevo país en paz”
Universidad de Pamplona
Pamplona - Norte de Santander - Colombia
Tels: (7) 5685303 - 5685304 - 5685305 - Fax: 5682750
www.unipamplona.edu.co