

**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

TÍTULO:

**Aplicación de Machine Learning para la clasificación, suscripción y
abandono/retiro de la suscripción de la Compañía Anónima el
Universo.**

AUTORES:

Astudillo Muñoz Juan Camilo

Villares Zambrano Jhon Michael

Trabajo de integración curricular previo a la obtención del título de

LICENCIADO EN NEGOCIOS INTERNACIONALES

TUTOR:

Ing. Carrera Buri, Félix Miguel, Mgs.

Guayaquil, Ecuador

7 de febrero del 2023

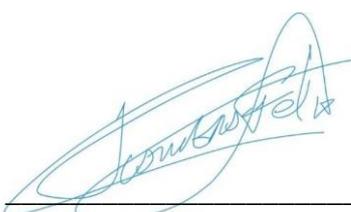


UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo de integración curricular, fue realizado en su totalidad por **Astudillo Muñoz Juan Camilo y Villares Zambrano Jhon Michael**, como requerimiento para la obtención del título de **Licenciado en Negocios Internacionales**.

TUTOR

f. 

Ing. Carrera Buri, Félix Miguel Mgs.

DIRECTOR DE LA CARRERA

f. 

Ing. Hurtado Cevallos, Gabriela Elizabeth Mgs.

Guayaquil, a los 7 del mes de febrero del año 2023



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Astudillo Muñoz Juan Camilo y Villares Zambrano Jhon Michael

DECLARAMOS QUE:

El Trabajo de Integración Curricular: **Aplicación de Machine Learning para la clasificación, suscripción y abandono/retiro de la suscripción de la Compañía Anónima el Universo**, previo a la obtención del título de **licenciados en negocios internacionales** ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de nuestra total autoría.

En virtud de esta declaración, nos responsabilizamos del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 7 días del mes de febrero del año 2023

LOS AUTORES:

f.

Astudillo Muñoz Juan Camilo

f.

Villares Zambrano Jhon Michael



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

AUTORIZACIÓN

Nosotros: Astudillo Muñoz Juan Camilo y Villares Zambrano Jhon Michael

Autorizamos a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Integración Curricular: **Aplicación de Machine Learning para la clasificación, suscripción y abandono/retiro de la suscripción de la Compañía Anónima el Universo**, cuyo contenido, ideas y criterios son de nuestra exclusiva responsabilidad y total autoría.

Guayaquil, a los 7 días del mes de febrero del año 2023

LOS AUTORES:

f. _____

Astudillo Muñoz Juan Camilo

f. _____

Villares Zambrano Jhon Michael

AGRADECIMIENTO

Quiero agradecer principalmente a las dos personas que han sido mi apoyo incondicional durante toda mi vida y más en este momento tan importante como lo es la culminación de mi carrera universitaria y un gran paso en mi vida profesional, mis padres Jorge Astudillo y Alejandra Muñoz. Sin el esfuerzo y sacrificio de ambos, esto no habría sido posible. Sé que he contado con el apoyo de ambos y estoy seguro de que seguiré contando con su apoyo el resto de mi vida. Por eso y muchas otras razones les dejo plasmado en este documento que, sin ustedes, mi vida no tendría sentido.

Agradecer a mis compañeros de universidad, como Joselyn, Rodrigo, Martha, Nicole, Mishelle, Jafet, y demás personas que he conocido a lo largo de este trayecto, como Génesis que sin ser de mi carrera se convirtió en una amiga muy importante para mí, personas las cuales han hecho este trayecto más ameno, llenando el transcurso de estos años de risas, enojos, peleas, tristezas y viaje, gracias a ellos la universidad será algo inolvidable.

A Leslie, que me ha acompañado a lo largo del último año, y ha hecho de este, el mejor que me haya podido imaginar. Gracias por ser incondicionalmente un apoyo y por motivarme siempre a ser mejor.

Por último, pero no menos importante, agradecer a mis profesores, a todos, por haber sido formadores de la persona que soy en este momento, pero principalmente a los que cruzaron la línea de ser profesores y se los puede considerar amigos. Gracias por haber sido parte de esta grata experiencia, estoy seguro de que los recordaré siempre.

-Juan Camilo Astudillo Muñoz

AGRADECIMIENTO

Agradezco a mis padres Stalin Villares y Elena Zambrano por siempre estar presentes en mi vida, por siempre alumbrarme el camino para poder seguir adelante en busca de cumplir todos mis objetivos y metas, y siempre brindarme consejo y aliento para continuar adelante.

Por último mi más grande agradecimiento a mis compañeros Angeline, Emily, Javier, Georgina, quienes me apoyaron en esta etapa educativa ayudándome a mejorar y poder dar el máximo de mí, también quiero agradecer a Raysha Macias por estar siempre para mí y brindarme su apoyo de manera incondicional durante todo este periodo.

Mi profundo agradecimiento a todas las autoridades y personal que hacen la Universidad Católica Santiago de Guayaquil, por brindarme la bonita experiencia de pertenecer a esta honorable institución, y agradecer a en especial a mis profesores quienes en mi han sido guías durante todo este camino educativo. Al ingeniero Félix Carrera quien ha sido el principal colaborador durante este proceso y quien con su dirección y enseñanzas permitió la creación de este trabajo.

-Jhon Michael Villares Zambrano

DEDICATORIA

Este logro va dedicado a mis padres, que son las principales personas que lo hicieron posible, Jorge Astudillo y Alejandra Muñoz. Estoy orgulloso de llevar estos apellidos acompañándome mi nombre, pues son estos los que me representan, en ellos llevo marcadas las palabras esfuerzo, sacrificio, valentía, coraje, amor, honestidad, educación, amabilidad y alegría. Los amo con todo mi corazón.

Este proyecto también va dedicado a mi hermanito, Jorge Andrés Astudillo Muñoz. Espero siempre ser el ejemplo al que apuntes seguir, y que esta tesis te demuestre que, con esfuerzo y dedicación, todo se cumple. Sueña en grande hermanito, que nuestros padres y yo siempre estaremos para apoyarte.

-Juan Camilo Astudillo Muñoz

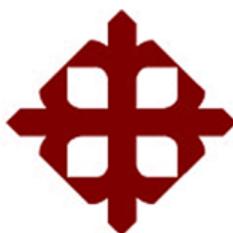
DEDICATORIA

El presente trabajo está dedicado a mis padres Jorge Stalin Villares Acurio y Carmen Elena Zambrano Brito, quienes con su apoyo han sido parte fundamental del motor que me llevo a cumplir un meta más en mi vida, gracias por brindarme un buen ejemplo de valentía, para no temerle a los problemas y adversidades que se me han presentado a lo largo de estos años y poder lograr cumplir todos mis objetivos y metas que me proponga.

A mis abuelos y tíos, quienes también han formado parte de este proceso y siempre han estado ahí para brindarme un consejo y palabras de aliento que me hicieron una mejor persona y hace que de una u otra forma esos consejos me acompañan a la hora de cumplir mis metas

Finalmente quiero dedicar esto a todos mis compañeros que nunca se negaron a brindarme su apoyo, a extenderme la mano en momentos en los cuales pasaba por dificultades, y por cada momento que brindamos juntos, siempre estarán presentes en mi vida.

-Jhon Michael Villares Zambrano



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

TRIBUNAL DE SUSTENTACIÓN

f. 

Ing. Hurtado Cevallos, Gabriela Elizabeth Mgs.
DIRECTORA DE CARRERA

f. 

Ing. Paredes Alcívar Fernando Andrés Mgs.
COORDINADOR DEL ÁREA

f. 

Ec. Freire Quintero César Enrique Phd.
OPONENTE



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

CALIFICACIÓN

Astudillo Muñoz Juan Camilo

Villares Zambrano Jhon Michael

Índice General

<u>RESUMEN</u>	XVIII
<u>ABSTRACT</u>	XIX
<u>RÉSUMÉ</u>	XX
<u>Capítulo I</u>	2
<u>Introducción</u>	2
<u>Problemática dentro de la empresa</u>	6
<u>Justificación</u>	8
<u>Alcance</u>	8
<u>Objetivos</u>	9
<u>Objetivo General</u>	9
<u>Objetivos Específicos</u>	9
<u>Capítulo II</u>	10
<u>Marco Teórico</u>	10
<u>Machine Learning</u>	10
<u>Aprendizaje supervisado</u>	12
<u>Modelo de Clasificación</u>	13
<u>Aprendizaje no supervisado</u>	14
<u>Matriz de confusión</u>	17
<u>Árboles de decisión</u>	17
<u>Naive Bayes</u>	18
<u>Redes Neuronales Artificiales (ANN)</u>	19

<u>Máquina de vectores de soporte (SVM)</u>	20
<u>Clustering</u>	20
<u>Big Data</u>	20
<u>Data Science</u>	21
<u>Inteligencia Artificial</u>	21
<u>Data Mining</u>	21
<u>Deep learning</u>	22
<u>Árboles de decisión</u>	22
<u>Marco Referencial</u>	22
<u>Marco Conceptual</u>	24
<u>Suscripción</u>	24
<u>Suscripción Gratuita</u>	24
<u>Suscripción de Pago</u>	24
<u>Promoción</u>	24
<u>Punto de Venta</u>	25
<u>Servicio Postentrega</u>	25
<u>Rentabilidad</u>	25
<u>Perfil del cliente</u>	26
<u>Servicios</u>	26
<u>Marco Legal</u>	27
<u>Capítulo III</u>	28
<u>Metodología</u>	28

<u>Teorema de bayes</u>	28
<u>Árbol de decisión</u>	29
<u>Pasos para desarrollar</u>	31
<u>Resultados</u>	34
<u>Análisis de Resultados</u>	36
<u>Conclusiones</u>	37
<u>Recomendaciones</u>	38
<u>Bibliografía</u>	39
<u>Anexos</u>	43

Índice de Tablas

<u>Tabla 1. Diferencia entre aprendizaje supervisado y no supervisado</u>	15
---	----

Índice de Figuras

<u>Figura 1. Ejemplo de gráficos de modelos de clasificación</u>	13
<u>Figura 2. Aprendizaje supervisado pasos</u>	14
<u>Figura 3. Matrices de confusión</u>	17
<u>Figura 4. Árboles de decisión</u>	18
<u>Figura 5. Naive Bayes</u>	19
<u>Figura 6. Redes Neuronales Artificiales</u>	19
<u>Figura 7. Máquina de Vectores de Soporte</u>	20
<u>Figura 8. Ejemplo de Árbol de Decisión</u>	30
<u>Figura 9. Árbol de Decisión Generado en RStudio</u>	34
<u>Figura 10. Forma Gráfica del Árbol Generado por RStudio</u>	35
<u>Figura 11. Matriz de Confusión</u>	35

ÍNDICE DE ANEXO

<u>Anexo 1. <i>Script completo en RStudio</i>.....</u>	43
<u>Anexo 2. <i>Base de Datos con Probabilidad y Predicción</i>.....</u>	44

RESUMEN

La capacidad humana de crear cosas que incrementen la aptitud de los individuos crece constantemente, es por esto por lo que la tecnología da pasos gigantes y, corresponde a las personas, seguirle el paso y adaptarse, ya que cada año surgen nuevas herramientas con las que es posible explotar un poco más el talento que se posee. El machine learning, inteligencia de negocios, inteligencia artificial y distintos con las que se conoce esta herramienta, es una de estas herramientas que logra aumentar la seguridad al momento de realizar una decisión empresarial, dando una idea de lo que representaría el tomar dicha decisión, esto gracias a que da un resultado certero, disminuye el error al fallo y, acompañado de una buena implementación es una herramienta de enorme valor en las decisiones empresariales.

Este proyecto tiene como finalidad demostrar que las herramientas como el machine learning pueden ayudar a prevenir situaciones que no se quieren en una empresa. Usando como ejemplo los suscriptores de la reconocida empresa C. A. El Universo. Para esto se usarán los modelos árbol de decisión, bosque aleatorio los cuales servirán para clasificar de forma correcta los suscriptores de esta empresa y evaluar la fidelidad de estos a dicha suscripción, según los resultados la empresa puede optar por una estrategia individual para cada persona y así, mantener al cliente suscrito al servicio, evitando una pérdida.

Palabras clave: Inteligencia de Negocios, Machine Learning, Árbol de Decisión, Bosque Aleatorio, Suscripción de Servicio.

ABSTRACT

The human capacity to create things that increase the fitness of individuals is constantly growing, which is why technology takes giant steps and it is up to people to keep up and adapt, since every year new tools emerge with which it is possible to exploit a little more the talent that one possesses. Machine learning business intelligence, artificial intelligence (and other names with which this tool is known) are some of the tools that manage to increase security when making a business decision, giving an idea of what it would represent to make said decision. This is thanks to the fact that it gives an accurate result and reduces the error or failure; and accompanied by a good implementation, it's of enormous value in business decisions.

The purpose of this project is to demonstrate that tools such as machine learning can help prevent situations that are not wanted in a company. Using as an example the subscribers of the renowned company C. A. El Universo. For this, the decision tree and random forest models will be used, which will serve to correctly classify the subscribers of this company and evaluate their fidelity to said subscription, according to the results the company can opt for an individual strategy for each person and thus, keeping the client subscribed to the service, avoiding a loss.

Keywords: Business Intelligence, Machine Learning, Decision Tree, Random Forest, Service Subscription.

RÉSUMÉ

La capacité humaine à créer des choses qui augmentent la forme physique des individus ne cesse de croître, c'est pourquoi la technologie fait des pas de géant et c'est aux gens de suivre et de s'adapter, car chaque année de nouveaux outils émergent avec lesquels il est possible d'exploiter un peu plus le talent que l'on possède. L'apprentissage automatique, l'intelligence d'affaires, l'intelligence artificielle et d'autres avec lesquels cet outil est connu, est l'un de ces outils qui parvient à augmenter la sécurité lors de la prise d'une décision commerciale, donnant une idée de ce que cela représenterait pour prendre ladite décision. ceci grâce au fait qu'il donne un résultat précis, réduit l'erreur à l'échec et accompagné d'une bonne mise en œuvre, est un outil d'une valeur énorme dans les décisions commerciales.

L'objectif de ce projet est de démontrer que des outils tels que l'apprentissage automatique peuvent aider à prévenir des situations indésirables dans une entreprise. En prenant comme exemple les abonnés de la célèbre société C. A. El Universo. Pour cela, les modèles d'arbre de décision et de forêt aléatoire seront utilisés, qui serviront à classer correctement les abonnés de cette entreprise et à évaluer leur fidélité audit abonnement, en fonction des résultats, l'entreprise peut opter pour une stratégie individuelle pour chaque personne et ainsi, en gardant le client abonné au service, en évitant une perte.

Mots clés: Business Intelligence, Machine Learning, Arbre de décision, Random Forest, Service Subscription.

Capítulo I

Introducción.

Actualmente estamos viviendo una nueva era en la que los negocios tienden a relacionarse de una manera distinta con sus clientes, ya que estamos presenciando un cambio sobre el manejo de la información de estos, las principales causas son la innovación tecnológica, la transformación de las necesidades y expectativas del cliente, por mencionar algunas. Por este motivo es indispensable para las empresas una transformación digital que nos permita competir en el mercado y entrar en un entorno económico más estable para poder sobrevivir en el futuro.

En este momento el uso de la inteligencia artificial ha tenido un gran apogeo dentro de este campo empresarial ya que con una recolección de datos eficaz y un software conveniente para el estudio se puede lograr tener una gran precisión de los posibles resultados que una empresa puede llegar a presenciar, ayudándonos a tener una mejor toma de decisiones y darle una mejor vía de desarrollo a esta.

Dentro de la inteligencia artificial podemos encontrar ramificaciones como lo puede ser el machine learning. Mediante conjuntos de algoritmos les brinda a los ordenadores la capacidad de identificar patrones para así darnos predicciones, estas predicciones estarían basadas en un conjunto de datos, mientras esta data tenga mayor extensión el ordenador nos proveerá con una predicción de mayor eficacia. Esta es empleada en cualquiera escala en la que se encuentre su empresa, pequeñas, medianas o grandes. (Anonimo, 2021)

Las empresas deben estar en constante actualización, y este es uno de los caminos que en estos últimos años ha logrado el mejor desempeño dentro de campos como marketing, logística, por citar algunos. Cada vez se va observando el aumento del ejercicio de estas tecnologías debido a la necesidad de lograr un futuro duradero dentro de los mercados, por lo que en unos años el mundo dejara atrás la manera convencional de tomar decisiones. (Anonimo, 2021)

Si bien la utilización de esta tecnología no ha sido recibida del mismo modo en las diferentes partes del mundo, hoy en día las empresas alrededor de todo el mundo han empleado este proceso, en ciertos lugares de manera más avanzada y en otros con un poco de retraso en cuanto a este salto a lo digital.

El continente europeo países como Alemania, Francia, España y Reino Unido el 87% de las más grandes empresas operan con estas tecnologías de Inteligencia artificial, de estas un 35% cuentan con un desarrollo completo, ayudándolos a escalar en esta transformación digital, pero demostrando que en el continente europeo a pesar de tener un buen porcentaje de desarrollo en empresas aún tiene camino por recorrer en cuanto al empleo de estas tecnologías. (Garay, 19)

Por otra parte, en Oceanía es claro que, en países como Nueva Zelanda y Australia, están adelantados con respecto a otros países pertenecientes al Continente, aquí el desarrollo de la inteligencia artificial no ha tenido un desarrollo masivo como podemos notarlos en el resto de los continentes. (Unesco, 2019)

En el continente asiático tenemos un giro drástico, ya que China se ha enfocado de manera específica en ser líder de sector de Inteligencia Artificial para el año 2030 compitiendo casi al mismo nivel que Estados Unidos, actualmente las mayores 20 empresas que se dedican a este sector son originarias de estos dos países, siendo 12 empresas americanas y las 8 restantes perteneciente a China. (Micó, 2018)

China está brindando un financiamiento cada vez mayor a empresas que formen parte de este grupo empleador de Inteligencia Artificial sin importar el área de acción al que pertenezcan.

Sanidad posee un mayor empleo, con un 19%, Múltiples industrias de un 17% empatando con la industria automovilística, Sin embargo, en empresas solo posee un 8%, aun así, podemos notar el gran impacto de estas empresas a escala mundial. Otros países como Japón y Corea del Sur tratan de seguir el paso de china, pero están un poco atrasados con respecto a este gigante de la inteligencia artificial.

En el continente africano si existe un gran retraso con respecto al empleo de esta tecnología, pero ciertos startups están apostando a inteligencia artificial, ciencia de datos, Deep learning ocupándolas en ciertos sectores, el sector agrícola es el más desarrollado, analizando tipos de cultivo mediante drones para obtener mejor eficiencia en la producción en ciertos sectores agrícolas. En el sector de seguridad se están implementando el uso de drones para el escaneo de millas en el océano para detectar barcos que puedan representar un riesgo ambiental, o que se dediquen a actividades ilegales, como podría ser la pesca ilegal. (Vishwana, 2021)

A pesar de esto apuntan a que se sigan desarrollando proyectos como estos de la mano de la inteligencia artificial en otros sectores en los que se encuentran grandes

problemáticas en el continente, sectores como salud no ha tenido grandes avances a pesar de lo suscitado hace unos años con el COVID 19, una de las causas de esto es que estos países no cuentan con los recursos necesarios ni una buena estrategia de manejo de esto para poder dar un salto abrupto al uso de estos sistemas, falta mucho para que en este lado del mundo compitan al igual que otros continentes en esta área de tecnología.

En el continente americano existe actualmente dos realidades, por un lado, tenemos ciertos países de América del norte, como son Canadá, Estados Unidos, que como mencionamos antes este último es una potencia en estas tecnologías, empleándolas en diferentes industrias, destacando por su manejo ciertos casos en sectores como atención al cliente, marketing para ventas. Al igual que China en el continente asiático Estados Unidos también apunta en ser el líder en este campo durante los siguientes años, promoviendo el desarrollo mediante inversión a la inteligencia artificial, el gasto ha ido aumentando constantemente desde el 2021 hasta el 2025, este incremento se estima en 120000 millones de dólares, representando un crecimiento del 26% anual, a nivel global este gasto representa el 50% de la inversión en inteligencia artificial a nivel global. (Álvarez, 2022)

Latinoamérica a pesar de tener la capacidad de aprovechar el sin número de beneficios que brinda la inteligencia artificial, debido a las diferentes limitaciones sociales y económicas que tienen los países que integran este bloque no se ha logrado un avance en esta área de tecnología, no existe un apoyo real por parte de los gobiernos para la integración a las industrias y la investigación de nuevas estrategias de desarrollo de estas. (Bielschowsky, 2019)

Esto es una desventaja para este bloque, ya que, dado que estas tecnologías pueden emplearse en múltiples áreas de acción, les brindaría mayor facilidad para poder contrarrestar el sin número de retos que están afectando al crecimiento y evolución de este bloque. (Bielschowsky, 2019)

Este bloque debe entender que estos avances no van a dar marcha atrás y su siguiente punto de alcance será una cuarta revolución industrial, a pesar de su situación actual estos deberían enfocarse en los atributos que le brindaría formar parte de este avance tecnológico si no se quiere quedar atrás en la evolución global que esta va desatando desde los últimos años.

Ecuador pertenece al bloque de Latinoamérica el cual no tiene grandes avances dentro de esta carrera, si bien está aumentando su uso, lo está haciendo sin un rumbo fijo.

Existen algunas causas para esto como la opinión pública que se tiene de estas tecnologías ya que no se ha proporcionado los beneficios y los riesgos de su empleo, también por este mismo desconocimiento a nivel legislativo no se ha logrado determinar un rumbo para la inteligencia artificial en el país.

Su uso se ha dado en gran parte en el sector privado, destacando los bancos, empresas de seguros y también en el área de salud con industrias farmacéuticas, su utilización destaca más que nada en contacto con los clientes, gestión, segmentación, verificar prospectos de clientes y evitar mora por parte de estos, como es el caso de instituciones como el Banco de Guayaquil, Produbanco. (Albornoz, 2020)

También en las universidades del país se están promoviendo proyectos basados en inteligencia artificial, es importante destacar que existe un buen ambiente para el desarrollo de este sector, ya que no se han registrado patentes por lo que no hay necesidad de pagar por derechos intelectuales, esto permite abaratar más el desarrollo de estos procesos.

El sector público también arranco a paso acelerado con estos sistemas en los últimos años, se han localizado ciertos problemas a los que se les ha otorgado una respuesta utilizando estas tecnologías, el hecho más notable fue en el 2020 por la pandemia de COVID 19, el IESS haciendo uso de un sistema de reconocimiento de lesiones en pacientes con COVID, permitió que se puedan realizar diagnósticos más eficaces basándose en tomografía de los afectados.

Aun así, con respecto a países vecinos como Colombia, Perú, Bolivia. Durante la pandemia Ecuador fue pionero en el empleo de inteligencia artificial, aun así, queda mucho camino por recorrer para avanzar en estos temas tecnológicos. Planificar una estrategia de Inteligencia artificial es primordial para un buen desarrollo de esta para poder beneficiar al país con oportunidades sociales y económicas.

Problemática dentro de la empresa.

La Compañía Anónima el Universo viene arrastrando inconvenientes hace años, cuenta con una cantidad bastante grande de suscriptores. Poco a poco la empresa logró tener una cartera de clientes/suscriptores bastante considerables, ya que la empresa cuenta con un increíble recorrido de más de 100 años en el mercado, de los cuales los últimos **20 años** se implementó un sistema de suscripciones. El cual consiste en el pago por llevar el servicio que ofrece la empresa al mercado hasta el hogar.

Es aquí donde la empresa comienza a tomar pasos que los acercan más hacia los clientes, en lugar de distribuir el servicio por revendedores como se hacía al inicio, la empresa buscó tener ese acercamiento que hace que los clientes se sientan especiales y por consecuencia con el paso de los años se obtuvieron suscriptores fieles los cuales cuentan con trayectos ininterrumpidos de renovar mensualmente su suscripción al servicio que ofrece la compañía por años, por lo que la empresa siempre busca en mejorar su servicio.

Sin embargo, la actual situación económica de la empresa la ha llevado a tomar medidas que permitan su sostenibilidad; reducir personal, reducir espacios de trabajo, maquinarias, etc. Es por esto, que la empresa se encuentra en busca de un método el cual le permita ahorrar, disminuir gastos y costos en la mayor cantidad posible. Es aquí donde entra la idea de prevenir pérdidas debido a los clientes que al final no renovarían la suscripción. La empresa realiza campañas de regalos hacia los suscriptores que posean las suscripciones más altas; tazas de regalos, entradas a conciertos, descuentos en distintas tiendas, etc. Estos regalos significan gastos que la empresa asume para mantener la fidelidad de la suscripción del cliente.

Para evitar cubrir gastos por clientes que a la larga no mantendrán una trayectoria de suscripción, se quiere emplear un sistema dinámico que use conceptos de machine learning el cual pueda predecir en base a la información de los clientes de la empresa si el nuevo suscriptor será responsable en los pagos y fiel al servicio, de esta forma se clasificará correctamente el nuevo suscriptor reduciendo el riesgo de pérdida. Además, se busca que el sistema cuente con un modelo de georreferenciación el cual ayudará a ubicar correctamente el lugar donde el servicio debe ser entregado, debido a que por direcciones muchas veces se complica la logística del reparto.

Actualmente la empresa no cuenta con un sistema que ayude a la toma de decisiones de los gerentes y analistas en distintos campos. Los sistemas implementados en la

compañía anónima el universo son programas que ayudan a la recolección de datos, procesos de compra, de recepción, de despacho, programas que facilitan el trabajo de los empleados. Sin embargo, carecen de un sistema que procese la información obtenida -la cual es valiosa- y gracias a los resultados faciliten en gran escala el proceso de toma de decisiones, ahorrando tiempo, liquidez y por lo tanto ayudando a mejorar con proyectos y estrategias dirigidas al mercado del que se es partícipe.

El hecho de no tener un sistema novedoso y eficaz ha dejado en evidencia los retrasos en la logística del reparto de las ediciones para los suscriptores, la dificultad de tomar las decisiones correctas en las reuniones de proyectos debido a que no tienen los suscriptores segmentados y agrupados dependiendo de sus características para poder dar a las estrategias un enfoque más directo. Es evidente que entre más grande sea una empresa, más difícil será mantener el orden y el control de esta, por lo que se ha presenciado distintos inconvenientes entre repartidores y sus rutas, retrasando el proceso de entrega y dando un mal servicio a los suscriptores.

Es por esto por lo que, el sistema que se busca implementar cuenta con una herramienta la cual es muy utilizada en las empresas del primer mundo. El machine learning, esta herramienta poco conocida en países de la Latinoamérica ayuda en gran medida y es de uso primordial en los negocios de países desarrollados, debido a que facilita la toma de decisiones mediante procesos estadísticos los cuales son de suma importancia para disminuir los errores en las decisiones tomadas, y por consiguiente obtener resultados que favorezcan en gran medida al máximo rendimiento de la empresa.

Justificación

La presente investigación se va a enfocar en analizar e identificar un patrón en el comportamiento de los suscriptores del diario el Universo. Esta investigación se debe a la reacción en el número de suscripciones vendidas, pero no renovadas. Tomando en cuenta que representa pérdidas económicas para la empresa que ciertos clientes no renueven la suscripción cuando tomaran los beneficios brindados por la empresa en ciertos meses en específico. También se puede agregar que uno de los beneficios de la suscripción es la entrega puerta a puerta, por lo que se realiza un mapa para el recorrido de los repartidores, por lo que clientes que se suscriben por un corto tiempo solo generan una mayor carga a los recorridos y un estudio individual para agregar la entrega del periódico a estos nuevos clientes en uno de estos recorridos. Este problema se está suscitando por el desinterés de los clientes en los beneficios que se les brinda mensualmente, tomando en cuenta de que estos cambian periódicamente por lo que es necesario un estudio que nos brinde estrategias para conseguir un mejor índice de renovación de suscripciones mediante la correcta clasificación de estos, dependiendo de ciertas variables intrínsecas, dentro del ámbito del estudio.

Alcance

El siguiente proyecto tiene como alcance a estudiantes de estadística, business intelligence y a la alta gerencia estratégica. Dentro de la actualidad los académicos deben comprender la importancia del uso de la inteligencia artificial orientada a los negocios, ya que el mundo en el que vivimos y el que se aproxima es un mundo tecnológico en su máxima expresión, en donde las personas que dominen esta clase de herramientas serán las personas que estarán aptas para competir en el mundo laboral. A las personas que quieren explotar el máximo rendimiento de sus empresas existentes o aquellas que, por temor a que las cosas no salgan bien no han tenido el valor de formar sus propios emprendimientos. A la alta gerencia estratégica se confía que el presente proyecto logrará que abran las puertas a una mejor forma de usar la estadística para la toma de decisiones empresariales, ya que la empresa cuenta con un recorrido bastante largo y han llegado al punto de no tomar riesgos en cuanto a las nuevas técnicas que se van dando con el paso de los años. La estadística descriptiva que usa la empresa hoy en día se considera obsoleta por lo que se prevé que el uso de modelos proyectivos o clasificatorios ayudará a tomar mejores decisiones a futuro.

Objetivos

Objetivo General

Realizar un modelo de machine learning para clasificar suscriptores y tomar decisiones estratégicas sobre la predicción del comportamiento futuro de los nuevos suscriptores.

Objetivos Específicos

- Analizar los conceptos y el marco teórico mediante la revisión literaria acerca del tema de investigación.
- Comprender la metodología de análisis supervisado para la aplicación de conceptos de machine learning.
- Describir los resultados sobre el análisis de datos y la ejecución del modelo.

Capítulo II

Marco Teórico

La inteligencia artificial, un tema que ha dado relevancia desde sus inicios en mediados del Siglo XX hasta la actualidad, siendo un tema extenso y complicado de estudiar y analizar, pues utilizan las ciencias exactas para su ejecución. El ser humano está en la constante búsqueda de buscar una herramienta que pueda generar o replicar los procesos cognitivos, pero de una manera que implique menos contacto humano. Según ha definido Rouhiainen (2018), la inteligencia artificial es la capacidad que poseen computadores, máquinas, entre otros de usar algoritmos, ya que pueden aprender de los datos que se obtienen para presentar resultados y tomar decisiones.

Es una ciencia que se viene desarrollando desde hace décadas desde la presentación de computadoras con índole de investigación científica en 1956. El desarrollo de esta ciencia automatiza los procesos humanos con la finalidad de la presentación de resultados, pero con un margen de error más pequeño. Villani(2018) manifestó “que existieron numerosos y variopintos métodos de la inteligencia artificial que han sido utilizados hasta la actualidad (ontológicos, aprendizaje por refuerzo, aprendizaje por confrontación, redes neuronales, etc.)” siendo estos partes de aplicaciones de resolución de problemas de ámbito empresarial, presentarse de manera informativa o de automatización de procesos.

Además, Rouhiainen (2018), manifiesta que la IA ya ha estado siendo utilizada por los seres humanos para lograr obtener eficiencia en varios ámbitos. Siendo el reconocimiento facial, protección cibernética, texto predictivo las más conocidas y comunes. Así, en el largo plazo, dentro del punto de vista empresarial, la IA proveerá ventajas competitivas para su crecimiento y predecir situaciones en las que pueden tomar decisiones prácticas y mejores en base a la recolección de datos de manera histórica.

Machine Learning

Dentro de los conceptos de la extensa ciencia de datos, una rama de la inteligencia artificial es el machine learning o aprendizaje automático, es aquella que tiene capacidad de aprendizaje y produce resultados sin necesidad de estar programados.

Desde la salida de computadores de casa y portátiles, simplificaron la vida de los consumidores, automatizando algunas actividades para promover la eficiencia del trabajo, reduciendo la carga laboral. “Por el contrario, gracias al aprendizaje automático, muchos de los dispositivos que verás en el futuro obtendrán experiencia y conocimientos a partir de la forma en que son utilizados para poder ofrecer una experiencia al usuario personalizada” (Rouhiainen, 2018). Por lo general, dentro de los comandos establecidos por los investigadores o usuarios de los softwares especializados emiten los resultados. Dentro del aprendizaje automático tomarían en cuenta las predicciones o el histórico de resultados creados con los comandos anteriores.

“Los avances en ML han permitido el reciente surgimiento de sistemas inteligentes con capacidad cognitiva similar a la humana que penetran en nuestra vida comercial y personal y dan forma las interacciones en red en los mercados electrónicos de todas las formas imaginables.” (Janiesch et al., 2020) Esto ocurre principalmente a que las compañías aumentan la productividad para la toma de decisiones, compromiso, adaptando a las preferencias del usuario. La capacidad de estos sistemas para la resolución avanzada de problemas, generalmente denominada inteligencia artificial (IA), se basa en modelos analíticos que generan predicciones, reglas, respuestas, recomendaciones o resultados similares.

Janiesch et al. (2020) también expresa que, “en las últimas décadas, ML ha progresado enormemente en su campo, con grandes avances y son notorios en los sofisticados algoritmos de aprendizaje y sus técnicas de preprocesamiento.” El machine learning supera esas mismas limitaciones, por tanto, los sistemas o avances tecnológicos serán cada vez más comunes al alcance humano y a grandes corporaciones que necesitan para maximizar la producción y reducir los costos de ciertos procesos gracias a la reducción de intervención humana en los mismos. Por esta razón, los algoritmos de ML se han aplicado con éxito en muchas áreas, como la detección de fraudes, la calificación crediticia, el análisis de la próxima mejor oferta, el reconocimiento de voz e imágenes o el lenguaje natural.

Es innegable que la evolución de las TICs ha cambiado drásticamente nuestras vidas y hubo un sinnúmero de mejoras en nuestros trabajos del día a día. Sin embargo, sin la existencia o aplicación del aprendizaje automático, habría asignaciones o tareas que implicaría un mayor grado de dificultad o de uso del tiempo, haciendo menos eficiente la labor en el trabajo. En cuantiosos casos, han transformado nuestros dispositivos o

herramientas electrónicas en verdaderas extensiones cognitivas que provocan un cambio en la manera en la que cambiamos las interacciones en el día a día. “Las TICs lograron este objetivo al llenar el vacío entre la percepción humana, el lenguaje, el razonamiento y el modelo y los instrumentos artificiales.” (Bonaccorso, 2017)

Dentro del concepto de la Inteligencia Artificial, el Machine Learning tiene varios tipos de aprendizaje, algunos más utilizados que otros. Sin embargo, cumplen la función de proveer al investigador o científico de datos la información requerida acorde al tipo de elemento o concepto del machine learning, ya que, usando de diferentes tipos de algoritmos para la solución de diferentes situaciones en la que se emplean los datos. A esto se le incluye, el tamaño de la población, el número de variables y el tipo de modelo en el que se pueda emplear de mejor manera para su ejecución.

Los avances en la tecnología han permitido que el machine learning o el aprendizaje de máquinas evolucione creando sistemas inteligentes con habilidades cognitivas iguales a las humanas, “de esta forma a nivel mundial se puede aumentar de formas inigualables la toma de decisiones” (Shrestha et al., 2021, p.600).

El machine learning se lo conoce como “la capacidad de crear sistemas para que aprendan de los datos de una forma automática para automatizar procesos o solucionar tareas asociadas por medio de algoritmos” (Janiesch et al., 2021, p.685). Nilsson (1998) también señala que el machine learning “se centra en varias disciplinas como la inteligencia artificial, estadística, teorías de la información, neurobiología, psicología y la complejidad computacional” (p.3). Tal como lo afirma López (2009) un algoritmo es “una sucesión en orden de procesos o instrucciones para solucionar un problema” (p.21).

Aprendizaje supervisado.

Aprendizaje supervisado: se enseña al algoritmo cómo realizar su trabajo (Rojas, E. 2018), con un conjunto de datos clasificados bajo una cierta apreciación o idea para encontrar patrones que puedan aplicarse en un análisis (Mueller, J., & Massaron, L., 2016) y producir una salida que ya se conoce.

En este tipo de aprendizaje, sus algoritmos trabajan en base a datos etiquetados, intentando descubrir una función que, según sus variables, “les asigne una etiqueta de salida adecuada a un nuevo valor, en otras palabras, predice su valor de salida.” (Santos,

2021) Cualquier sistema de predicción puede verse dentro de esta clasificación. Un ejemplo claro del aprendizaje supervisado son los árboles de decisiones.

Entonces, se entiende que el algoritmo produce una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema y constituye un algoritmo de aprendizaje basado en ejemplos donde el nuevo conocimiento es inducido a partir de una serie de ejemplos y contraejemplos.

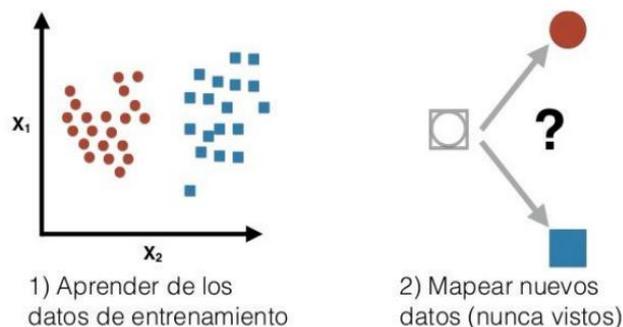
“Este tipo de aprendizaje sucede cuando pasa con un “maestro” el cual posee conocimientos del entorno” (Suresh et al., 2012, p.20). Este tiene dos secciones que son de datos de entrenamiento y datos de prueba, los de entrenamiento son usados para entrenar un modelo y los de prueba para ver la eficacia del modelo que se creó. El modelo que usa este tipo de aprendizaje es el Perceptrón.

Modelo de Clasificación.

Es categórico o su variable de salida es discreta. Esto quiere decir que existe un conjunto finito de posibles resultados de predicción. Esta se divide en binaria (0 o 1) y de multi-clase (múltiples categorías).

Figura 1.

Ejemplo de gráficos de modelos de clasificación



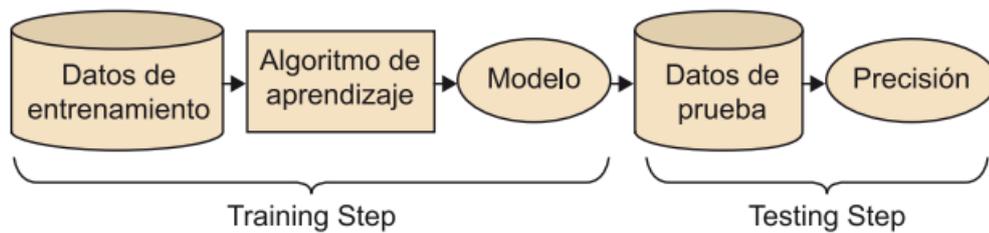
Nota. Tomado de Python Diario

Para el aprendizaje supervisado, “se usan algoritmos como: árboles de decisión, regresión por mínimos cuadrados, regresión logística, clasificación de Naïve Bayes o métodos ensemble” (Lu y Bai, 2010, p.326). Según Gironés (2016) afirma que “existen dos pasos importantes en el aprendizaje supervisado los cuales serían el entrenamiento y el paso de prueba” (p.8).

Como se puede observar en la figura, según ciertos datos de entrenamiento, que son patrones de entrada y salida, estos se acoplan a un modelo de aprendizaje, los cuales dan un modelo donde su nivel de precisión es testeado según datos de prueba. “Entonces, el sistema aprende un mapeo de una correcta salida para cada uno de los patrones de entrada” (Chapelle et al., 2006, p.2).

Figura 2.

Aprendizaje supervisado pasos



Nota. Tomado de Bing Liu. Web Data Mining.

Aprendizaje no supervisado

En este caso a diferencia del anterior los datos no están etiquetados. Por ende, lo único que se puede hacer es especificar la estructura de los datos para de esta forma hallar una forma de organizarlos, estos datos son de carácter exploratorio.

Como lo plantea el autor “en este aprendizaje, no es del todo necesario que los humanos sean parte de la preparación o elaboración de un conjunto de datos, un ejemplo de esto es el clustering” (Godoy, 2017, p.106). Madhulatha (2012) expresa que “hay dos tipos de algoritmos de clasificación no supervisada” (p.719):

Algoritmos jerárquicos (bottom-up). Los resultados se representan mediante dendogramas, estos se dividen en aglomerativos o bottom up y divisivos o top down.

Algoritmos particionales (top-down). En donde se tienen los nodos desde la optimización de una adecuada función para el objetivo del estudio.

Aprendizaje reforzado

Los autores Sutton y Barto (2015) plantean que el aprendizaje reforzado es “que cierto agente aprenda a hacer acciones que está en cierto estado”(p.2), es decir, este ayuda a perfeccionar la respuesta del modelo por medio de la retroalimentación. “Los

algoritmos de este tipo de aprendizaje normalmente están basados en la aproximación de una función de valor” (Recabal, 2009, p.12). Al contrario de los otros tipos que intentan reducir los errores aquí se maximizan las recompensas.

Diferencia entre aprendizaje supervisado y no supervisado

Tabla 1.

Diferencia entre aprendizaje supervisado y no supervisado

	Aprendizaje Supervisado	Aprendizaje No Supervisado
Método	Se debe dar variables de entrada y de salida.	Solo necesita datos de entrada.
Características	<ul style="list-style-type: none"> • Dispone de una característica objetivo para cada una de las instancias. • Modela relaciones y dependencias entre el resultado predictivo objetivo y las características de entrada. • Requiere mayor tiempo de entrenamiento. • Aprende de datos etiquetados. 	<ul style="list-style-type: none"> • No es necesario disponer de la respuesta correcta en los datos de entrenamiento. • La red descubre en los datos de entrada y de forma autónoma: características, regularidades, correlaciones y categorías. • Suelen requerir menores tiempos de entrenamiento. • No necesitan de un profesor externo.
Desventajas	<ul style="list-style-type: none"> • Tenemos que el proceso de entrenamiento suele 	<ul style="list-style-type: none"> • No hay un esquema general bien definido para

	<p>ser lento y no es infalible, se depende bastante de la elección de los casos de entrenamiento para que el clasificador sea capaz de generalizar lo suficiente.</p> <ul style="list-style-type: none"> • Es preciso un trabajo previo de clasificación manual de los casos que se usarán para el entrenamiento, que pueden ser muchos miles en un problema de cierta complejidad. 	<p>guiar su implementación.</p> <p>En cada problema tendrá una forma distinta y esto puede ser un inconveniente serio.</p>
Objetivo	Determinar una función para predecir una salida.	Determinar una función para predecir una salida.
Ejemplos	Clasificación, Asociación, Regresión Lineal.	Clustering, Asociación, K mean
Usos	Reconocimiento de imágenes, reconocimiento de voz, pronósticos, análisis financieros y capacitación de redes neuronales y árboles de decisión, etc.	Preprocesar los datos, durante el análisis exploratorio o pre-entrenar algoritmos de aprendizaje supervisado.

Técnicas de principales clasificaciones

Matriz de confusión

Esta es una herramienta utilizada para evaluar los modelos de clasificación. “Lo que comprende esta matriz es un grupo de valores que llevan la cuenta del grado de cuánto se asemejan las observaciones emparejadas, de un conjunto de bajo control (CDC) y uno de referencia (CDR)” (Ariza et al., 2018, p.216).

Figura 3.

Matrices de confusión

		Clase verdadera		
		C1	C2	
Clase predicha	C1	a	b	P_1
	C2	c	d	P_2
		π_1	π_2	

	Realmente es positivo	Realmente es negativo
Predicho como positivo	Verdaderos Positivos (VP)	Falsos Positivos (FP)
Predicho como negativo	Falsos Negativos (FN)	Verdaderos Negativos (VN)

Nota. Tomado del Libro “Algoritmos” Jordi Gironés.

Árboles de decisión

Según Zhao (2013) “cuando se trata de casos de estudio de predicciones estos normalmente son tratados por medio de árboles de decisión” (p.15). Sirven para representar sucesos que se originen a partir de una decisión de una forma analítica y gráfica, de esta manera permite ver y organizar un problema, para Echavarría et al., (2013) árboles de decisión “ayudan a reconocer entre las decisiones actuales y las futuras su respectiva interdependencia” (p.6).

De acuerdo con Wanjun y Xiaoguang (2010) afirman que “entre los algoritmos para hacer árboles de decisión están: SPRINT, SLIQ, PUBLIC, ID3, C4.5” (p.254). “Los elementos del árbol de decisión son cuatro, básicamente es su estructura” (Damodaran, 2007, p.25):

Nodo de decisión.

Dado por un cuadrado es el que indica que en dicho punto es necesario tomar una decisión.

Nodo de probabilidad.

Dado por un círculo da a conocer que en ese punto hay un evento aleatorio.

Rama.

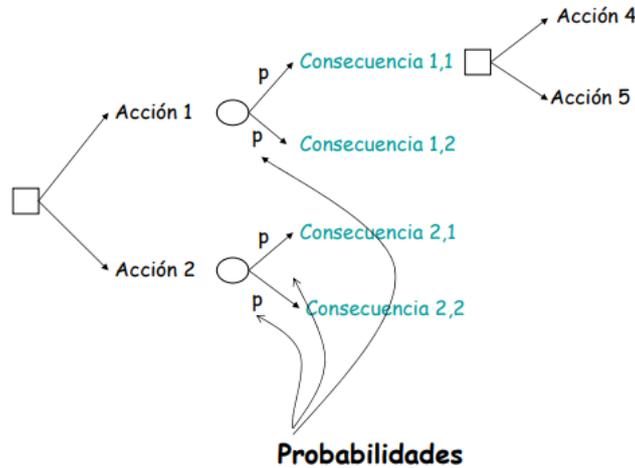
Son los caminos que llevan a tomar una decisión.

Nodo final.

Son los resultados.

Figura 4.

Árboles de decisión



Nota. Tomado de Damodaran “Probabilistic approaches”

Naive Bayes

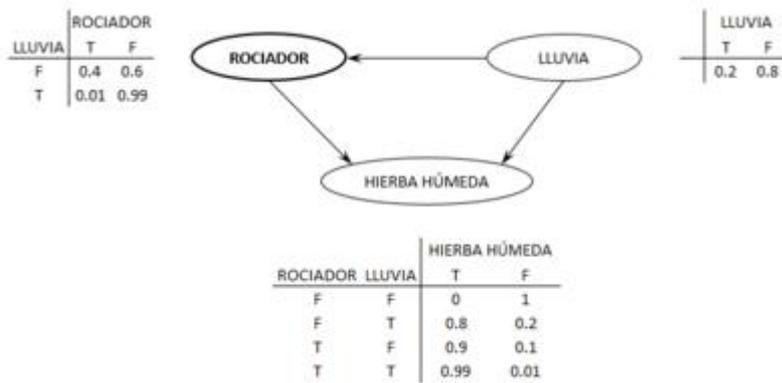
Las redes bayesianas son un método que da análisis de tipo cualitativo ya sea de los valores o de los atributos del problema, pero también ofrece la importancia cuantitativamente. “Son redes gráficas que carecen de ciclos por donde se pueden visualizar las relaciones de probabilidad entre ellas y las variables aleatorias, los nudos son las variables aleatorias y las relaciones entre ellas son los arcos” (Doldán, 2007, p.3).

El autor Mitchell (1997) menciona que “esta forma de clasificación es aplicada a las tareas de aprendizaje donde cada instancia x está descrita por un conjunto de valores atributos y donde la función objetivo $f(x)$ puede tomar cualquier valor de un set finito V ”(p.177). Rivera (2011) describe que “las redes bayesianas se dividen en continuas las cuales poseen un número infinito de valores posibles como la variación de la temperatura y las dinámicas que tienen una variable que es aleatoria por cada intervalo de tiempo” (p.3).

Nota. Tomado de Felix Doldán “Redes bayesianas y riesgo operacional”.

Figura 5.

Naive Bayes



Redes Neuronales Artificiales (ANN)

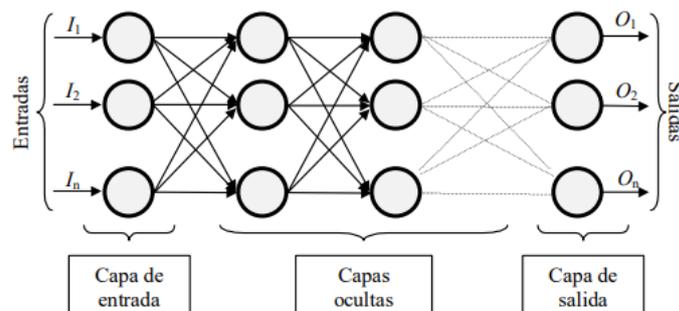
Gurney (1997) plantea que:

Una red neuronal es un conjunto que se encuentra interconectado de unidades, elementos, donde la capacidad de procesamiento de la red esta almacenada en las fuerzas de conexión entre unidades, los cuales son obtenidos de un transcurso de aprendizaje de un conjunto de patrones de entrenamiento. (p.13)

Por su constitución las redes neuronales artificiales muestran varias características semejantes a las del cerebro, “ofreciendo varias ventajas como el aprendizaje adaptativo, la auto-organización, tolerancia a fallos, la operación en tiempo real y la fácil inserción” (Matich, 2001, p.8). A continuación, se muestra en la figura un esquema de una red neuronal, la cual tiene neuronas interconectadas en tres capas, donde los datos entran por la capa de entrada, atraviesan por la capa oculta y salen por la capa de salida.

Figura 6.

Redes Neuronales Artificiales



Nota. Tomado de Damián Matich “Redes Neuronales: Conceptos Básicos y Aplicaciones.

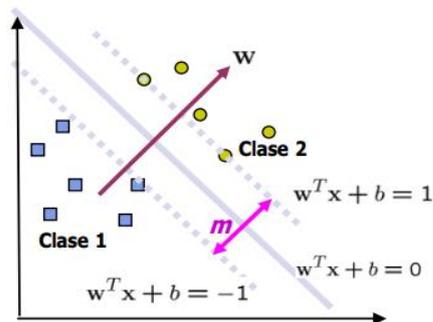
Estas se clasifican en binarias las cuales toman valores dentro del intervalo $\{0,1\}$ o $\{-1,1\}$ y las reales que van dentro de $[0,1]$ o $[-1,1]$.

Máquina de vectores de soporte (SVM)

Tal como lo afirma el autor Betancourt (2005) “esta técnica se basa en minimizar el riesgo estructural (SRM), donde los puntos de entrada son mapeados en un espacio donde las características tienen una dimensión mayor y de esta manera haya un hiperplano por el cual los separa para maximizar el margen m ” (p.67). En la siguiente imagen se puede observar que la frontera de decisión necesita estar alejado de los datos de las dos clases para tener una buena clasificación.

Figura 7.

Máquina de Vectores de Soporte



Nota. Tomado de Gustavo betancourt “Las máquinas de soporte vectorial”

Clustering

Este proceso de clasificación se usa en el aprendizaje no supervisado, se lo puede definir como “es un procedimiento multivariante repartir un conjunto de datos en un conjunto de subclases homogéneas los que se los denomina clusters” (Schiaffino, 2018, p.2). El método k-means es uno de los algoritmos principales de clustering, Morisette y Chartier (2013) establecen que “esta técnica es una aproximación que supone que los componentes de la mezcla poseen matrices de covarianza esféricas y probabilidades de muestreo iguales” (p.15). Es decir, de esta forma se puede obtener información sobre grandes conjuntos de datos multivariados, determinando variables y escogiendo los clusters.

Big Data

El Big Data se lo puede conocer como un activo de información de mucha variedad, volumen y velocidad, “el cual tiene datos estructurados y no estructurados, el

Big Data ayuda a ocupar varias ramas de la inteligencia artificial, como los modelos predictivos” (Salvador y Cárdenas, 2020, p.2).

Data Science

Este surge como “un método por medio del uso de tecnologías de Business Intelligence para de esta manera predecir, dar puntos de vista históricos y actuales, donde el Big Data tiene la función del análisis de la información, presentación de informes y la recopilación de datos” (Goyzueta, 2015, p.149).

Inteligencia Artificial

El machine learning es uno de los enfoques más importantes de la inteligencia artificial. Según Rouhiainen (2018) especifica que “la IA es la capacidad que poseen las máquinas para aprender de los datos, usar algoritmos y tomar decisiones de la forma en la que lo quería un humano” (p.17). Con la diferencia que este no necesita descanso y los errores en las tareas son menores.

La inteligencia artificial y su desarrollo actual tienen como componente núcleo, los datos, pese a que se utilizan la matemática y la estadística para su ejecución, también es imperativo la recolección y manejo de datos, así estos jugarán un rol importante en varios sectores de la producción. Desde el blockchain hasta los procesos super computacionales han requerido de los algoritmos o modelos que utilicen la IA para el desarrollo de un procedimiento que implique el ahorro del tiempo y un minúsculo margen de error. Hasta ese entonces Silicon Valley, era conocida como un hub tecnológico, debido a que las grandes corporaciones tecnológicas posicionaron sus oficinas centrales ahí, presentando innovaciones cada año, que revolucionan el estilo de vida de la humanidad.

Data Mining

“Es un proceso por el cual se obtienen modelos o patrones desde los datos que fueron recopilados, para hallar situaciones anómalas, padrones, tendencias o secuencias” (Valcárcel, 2004, p.84). Dentro de la data mining se encuentran tres componentes que son el modelo, el criterio de preferencia y el algoritmo de búsqueda. Donde se usan técnicas de verificación y métodos de descubrimiento.

Deep learning

Es un tipo de machine learning, que en vez de ordenar los datos para que se puedan ejecutar, realiza la configuración de ciertos parámetros de los datos y entrena a la máquina para que de esta manera pueda aprender automáticamente patrones. Actualmente se usa popularmente en reconocimiento del habla, de imágenes, sistemas de recomendaciones. Las redes neurales son unos de los componentes más importantes del Deep learning.

Árboles de decisión

Este algoritmo está basado en el teorema de Bayes, un árbol de decisión o diagrama es un modelo de la evaluación de una función discreta, en la que se determina el valor de una variable y la siguiente acción (elegir otra variable para evaluar o generar el valor de la función) se elige en consecuencia. (Moret, 1982) En otras palabras, los árboles de decisiones dentro del machines learning, ofrece al usuario una herramienta visual para predecir un hecho mediante la probabilidad y clasificación de datos.

Marco Referencial

Es necesario analizar el uso de estas prácticas para la clasificación de clientes en otras investigaciones que poseen ciertas características de importancia para este estudio, para poder realizar un correcto análisis de los resultados de la investigación. Por esto vamos a tomar de referencia algunas investigaciones que se han venido realizando en los últimos años.

Vamos a empezar con la investigación realizada por (Espinosa-Zúñiga, Scielo, 2020), la cual se basa en el desarrollo de un Random Forest (RF) para determinar si un postulante es o no es apto para recibir una tarjeta de crédito, basándose en la base de datos donada por una institución bancaria perteneciente a Australia para su estudio en la Universidad de California en 2019. Este estudio empezó con un análisis exploratorio previo al desarrollo del (RF), donde se pudieron destacar las principales características de dicha base de datos, entre las cuales están la cantidad de datos de la base, que sus columnas fueron modificadas desde la fuente con el fin de proteger la confidencialidad de esta, y que posee 16 columnas, 15 predictoras y una variable objetivo.

Con esto esta institución financiera puede realizar una correcta evaluación de sus postulantes, y por ende tomar una decisión más acertada con respecto a su viabilidad como acreedor de una tarjeta de crédito de esta institución.

En segundo lugar, vamos a mencionar la investigación realizada por (López, 2021) la cual se desarrolló para la empresa Insofec, empresa ecuatoriana ubicada en la Concordia, con el fin de obtener una clasificación de sus clientes mediante Random Forest, como “Buenos” o “Malos” pagadores, tarea que se estaba realizando de manera manual lo cual provocaba una carga excesiva de trabajo operativo para poder realizar dicha calificación, para esto la empresa brindo la base de datos histórica de sus clientes desde el 2017 hasta el 2020. Esta base cuenta con 18 variables y 63896 registros. Este número de registros disminuyo cuando se realizó la eliminación de datos atípicos, se procedió a realizar el modelo en Random Forest. Este modelo resulto con una precisión del 97%, y tan solo un 2.8% de error, con lo cual podemos decir que este modelo es lo suficientemente eficiente como para realizar una transformación en el sistema de calificación de clientes que se tenía en su momento en la empresa Insofec.

Como tercer punto de referencia vamos a mencionar una investigación que se realizó en Perú por (VILCA & CAMARGO RAMOS , 2021) la cual se basa en la deserción de clientes de una institución financiera, ya que no contaban con un índice de probabilidad de deserción de sus clientes para microcréditos en la banca de la ciudad de Lima, lo cual presenta un gran problema la esta institución, como primer paso se realizó un análisis a la base de datos, para posteriormente, realizar la eliminación de datos inconsistentes en esta. También se destacaron las principales variables a utilizar para esta predicción, las cuales serían la cantidad de transacciones promedio, y los montos promedio de consumo de los clientes. Para posteriormente realizar el modelo de Random Forest de la mano de una Red Neuronal y una Maquina de Soporte Vectorial para obtener una precisión mayor en este modelo, una vez se ingresaban la data al modelo pase por estos 3 filtros, para una vez finalizado mediante un Perceptrón Multicapa arrojar el resultado de si es un posible desertor o no.

Marco Conceptual

Suscripción

La suscripción es la acción que realiza un usuario cuando quiere recibir una serie de servicios, noticias, o productos a un precio determinado, o de forma gratuita. (Peiro, 2019). Como mencionamos antes existen suscripciones de pago o gratuitas, las cuales son utilizadas de manera más frecuente en cuestiones digitales, pero son distintas ya que poseen objetivos y características diferentes.

Suscripción Gratuita

Son aquellas que se llevan a cabo normalmente cuando un usuario se suscribe a un blog para recibir contenido gratuito por parte del especialista que lo escribe. El objetivo por parte del suscriptor es aprender, o mejorar su conocimiento sobre un tema que le interesa de forma gratuita (Peiro, 2019). Por otra parte, el dueño del blog está interesado en conseguir el correo electrónico de las personas para posteriormente elaborar campañas de email-marketing.

Suscripción de Pago

En este caso existe un importe a cambio de recibir información, cursos, o documentos. El suscriptor se suscribe para recibir contenido premium, o privado, ya que no está disponible para todo el público.

Una empresa puede disponer de esta opción para todos aquellos interesados que quieran ampliar conocimientos a través de la opción de recibir contenido novedoso de forma habitual a través de una suscripción mensual, por ejemplo. (Peiro, 2019)

Promoción

La promoción es "el conjunto de técnicas integradas en el plan anual de marketing para alcanzar objetivos específicos, a través de diferentes estímulos y de

acciones limitadas en el tiempo y en el espacio, orientadas a públicos determinados" (Bonta & Farber, 1994)

Pero en un sentido general, la promoción es una herramienta táctica-controlable de la mezcla o mix de mercadotecnia (4 p's) que combinada con las otras tres herramientas (producto, plaza y precio) genera una determinada respuesta en el mercado meta para las empresas, organizaciones o personas que la utilizan. (Thompson, 2005)

Punto de Venta

Un punto de venta, por tanto, es el lugar físico (tienda), o virtual (ecommerce), en el que una empresa tiene un contacto con un posible consumidor. Este espacio, como decíamos, puede ser físico, es decir, puede existir físicamente, pudiendo el cliente acudir a él. O, por otro lado, puede tratarse de un espacio virtual, un portal web, donde se halla un negocio online. (Morales, 2021)

Servicio Postentrega

El servicio post entrega es cualquier tipo de atención que se le proporciona a un cliente después de que haya comprado un producto o servicio. Este tipo de atendimento es muy utilizado por las empresas como estrategia comercial, ya que normalmente conduce a una mayor satisfacción del cliente, lealtad hacia la marca e incluso marketing de boca en boca. (Silva, 2020)

Rentabilidad

La rentabilidad hace referencia a los beneficios que se han obtenido o se pueden obtener de una inversión. Tanto en el ámbito de las inversiones como en el ámbito empresarial es un concepto muy importante porque es un buen indicador del desarrollo de una inversión y de la capacidad de la empresa para remunerar los recursos financieros utilizados. (Arias, 2015)

Perfil del cliente

El **perfil de cliente** es un conjunto de información demográfica, psicográfica y de comportamiento que caracteriza al comprador ideal de una marca. Su objetivo es facilitar la atracción, nutrición y retención de consumidores.

Definir un perfil de cliente para una empresa significa recopilar datos relevantes que la ayuden a crear mejores productos, servicios y ofertas para su público objetivo. De esa forma, será más fácil para el prospecto avanzar a través del embudo de conversión. (Silva, Zendesk, 202)

Servicios

Un servicio es una actividad o una serie de actividades llevadas a cabo por un proveedor con el propósito de satisfacer una determinada necesidad del cliente, sus características básicas son la intangibilidad, la heterogeneidad, la inseparabilidad y la perecibilidad. (Alfonsp, 2019)

Marco Legal

En el ámbito legal se han generado varios planes alineados a una normativa legal entre los cuales se destacan: El Plan Nacional de Telecomunicaciones y Tecnologías de Información del Ecuador 2016-2021. El Convenio Marco de Cooperación Interinstitucional firmado con los diferentes entes como el Ministerio de Telecomunicaciones y de La Sociedad de la Información, la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación.

Los Planes Nacionales para el Desarrollo (2013-2017). La Ley Orgánica de Telecomunicaciones. El Código Orgánico de La Economía Social de los Conocimientos, Creatividad e Innovación. Así como los artículos que contemplan el desarrollo de la investigación científica y de la innovación tecnológica, de las TIC y diversas áreas del conocimiento que están en la Constitución de la República del Ecuador de los cuales se puede destacar los (Art. 281, Art. 385, Art. 387, Art. 423, Art. 388, Art. 313) que permiten alinearse y ser partícipes de la adopción de estas tecnologías emergentes para el cambio y adopción de una cultura de transformación digital para un mejor desarrollo a nivel empresarial en el ámbito local con proyección internacional, lo que se convierte en un impacto positivo.

Capítulo III

Metodología

Para esta investigación se va a emplear los algoritmos de árbol de decisión y random forest para la clasificación. El árbol de decisión es un algoritmo que se puede utilizar tanto para la regresión como también para la clasificación, la base de este algoritmo es evaluar todas las variables de entrada y todos los puntos de división posibles para seleccionar el mejor de todos, la razón por la que se llama árbol de decisión es debido a su estructura que se asemeja a un árbol, pero al revés. En este algoritmo se divide la data en dos o más conjuntos homogéneos basados en el diferenciador más significativo en las variables de entrada. Este algoritmo determina la variable con más significancia y el valor que proporciona los mejores conjuntos homogéneos de la población. El segundo algoritmo que usaremos es un random forest, es llamado un método en conjunto, lo cual quiere decir que los resultados se combinan, pero ¿cuáles resultados?, básicamente de todos los árboles de decisión que componen el bosque. Cada árbol se entrena en un subconjunto de la serie de datos y da un resultado (sí o no). Luego de esto, se combinan los resultados de todos los árboles de decisión para dar una respuesta final. Cada árbol “vota” (sí o no) y la respuesta final es la que tenga la mayoría de los votos.

Teorema de bayes

El teorema de Bayes fue enunciado por Thomas Bayes, este se desarrolló para el cálculo de probabilidades, pero de una manera contraria a como se calculan comúnmente.

Normalmente se podría calcular la probabilidad de que suceda algo (A) sabiendo que este suceso está cumpliendo con algún tipo de característica que está condicionándolo. El teorema de Bayes entiende la probabilidad de forma inversa al teorema de la probabilidad total. El teorema de la probabilidad total hace inferencia sobre un suceso B, a partir de los resultados de los sucesos A. Por su parte, Bayes calcula la probabilidad de A condicionado a B. (López, 2018)

La fórmula del teorema de Bayes es $P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right) \cdot P(A)}{P(B)}$ donde podemos describir que B es el hecho que ya conocemos, A es el grupo de causas que podrían producir el hecho B, siempre excluyéndose una de la otra. Por lo tanto $P\left(\frac{A}{B}\right)$ y $P(A)$ son las posibilidades posteriores y anteriores respectivamente, dándonos que $P\left(\frac{B}{A}\right)$ es la probabilidad de que suceda b en cada característica condicionante de A.

Por ejemplo si viajamos en una motocicleta y esta empieza a detenerse, podemos obtener las probabilidades por las que esto puede suceder, esto puede ser que la motocicleta tuvo un fallo mecánico, o que se quedó sin combustible, podemos desarrollar esto un poco más y determinar que si circulan 11 motocicletas con diferentes características (4 son de alta gama y se encuentran al día en los mantenimientos, 3 son de gama media, y no se han saltado un par de mantenimientos, las otras 4 son de tercera mano y están totalmente descuidadas) y una de estas motocicletas se detiene por las razones antes mencionadas, también podemos calcular la probabilidad de a qué grupo pertenece la moto que tuvo el suceso.

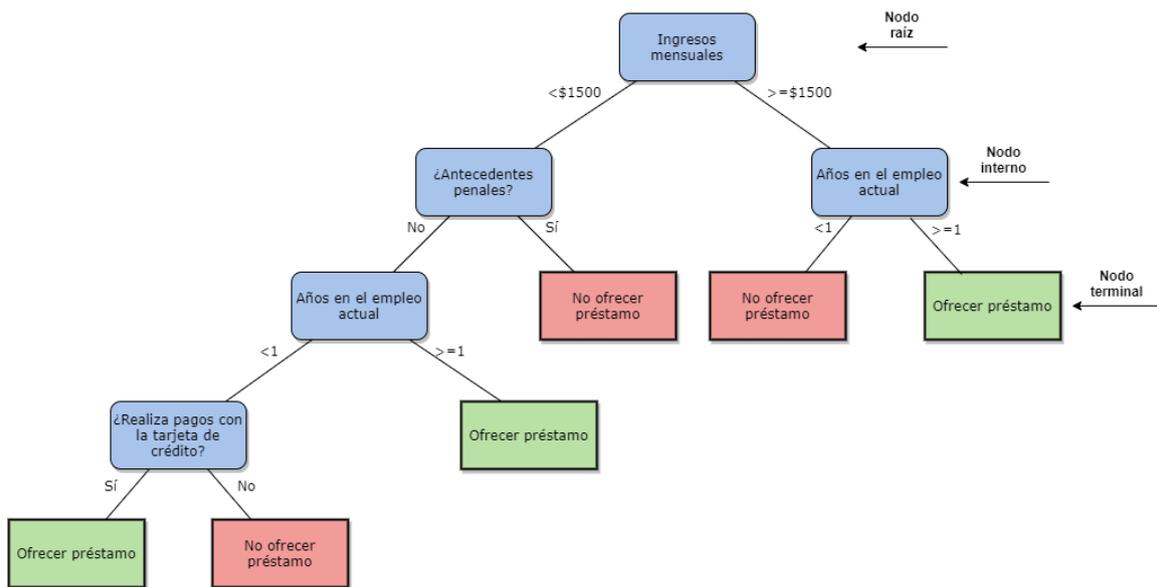
Árbol de decisión

Este algoritmo se puede utilizar tanto para la regresión como también para la clasificación, la base de este algoritmo es evaluar todas las variables de entrada y todos los puntos de división posibles para seleccionar el mejor de todos, la razón por la que se llama árbol de decisión es debido a su estructura que se asemeja a un árbol, pero al revés. En este algoritmo se divide la data en dos o más conjuntos homogéneos basados en el diferenciador más significativo en las variables de entrada. Este algoritmo determina la variable con más significancia y el valor que proporciona los mejores conjuntos homogéneos de la población.

Todas las variables de entrada y todos los puntos de división que sean posibles se evalúan y se elige la que nos brinde un mejor resultado. Los algoritmos de aprendizaje basados en árbol se consideran uno de los mejores y más utilizados de aprendizaje supervisado. Los métodos basados en arbole potencian modelos predictivos de alta precisión, estabilidad y facilidad de interpretación, a diferencia de los modelos lineales

mapean bastante bien las relaciones no lineales. El funcionamiento del algoritmo toma la variable más significativa se encuentra en la base del algoritmo creando un condición que nos brinda dos respuestas, una positiva y una negativa a partir de aquí se generan nuevas opciones y cada una de ellas genera nuevas opciones, siempre dando las mismas respuestas, positiva o negativa, las cantidad de condicione que se van a evaluar depende de la cantidad de variables independientes que encontremos, al final de realizar la evaluación de todas las características el algoritmo nos brinda una imagen en forma de árbol invertido.

Figura 8.
Ejemplo de Árbol de Decisión.



Como podemos observar la estructura puede dividirse en partes, comenzamos por el nodo raíz que es la primera condición de todo el algoritmo, es la base de todo el algoritmo, seguido tenemos los nodos internos o los nodos de pruebas, que son todas las condiciones y pruebas que se deben realizar para llegar al siguiente nodo, este es llamado nodo terminal o decisión, que sería el resultado o predicción del algoritmo, al ser un algoritmo de aprendizaje supervisado es necesario que se introduzca las variables dependiente e independientes, las variables independientes es la que se manipula para determinar el valor de una variable dependiente, es decir son las característica que queremos usar para predecir un valor de Y, la variable dependiente es la que depende de la variable independiente, en resumen esta es la variable que deseamos predecir.

Pasos para desarrollar

Para la creación del modelo vamos a emplear Rstudio, el cual nos brinda una interfaz, en la que se permite tener un mejor manejo y más fluidez con el lenguaje R por medio del programa Rstudio. Para poder realizar el modelo tenemos que llamar a la base de datos de los cual se encarga la función (read.csv) o (read.csv2) dependiendo del tipo de base de datos que llamemos se escogerá una u otra.

```
Suscriptores <- read.csv2("../Data tesis/Suscriptores 2.csv")
```

Una vez tengamos la base de datos lista podemos empezar a programar las funciones necesarias para poder crear el árbol, para esto necesitamos descargar una serie de paquete que se detallaran a continuación:

El primero es el paquete **Caret** (abreviatura de Clasificación y entrenamiento de regresión), la cual nos brinda funciones para agilizar el proceso de entrenamiento del modelo para problemas complejos de regresión y clasificación. También nos ofrece una interfaz unificada que simplifica el proceso de modelado empleando la mayoría de los métodos de AE implementados en R. Además de proporcionar rutinas para los principales pasos del proceso, incluye también numerosas funciones auxiliares que permitirían implementar nuevos procedimientos.

Otro paquete que se usa es el Rpart, en este paquete se encuentra lo necesario para el desarrollo de los árboles de clasificación que usaremos, es un paquete que le da forma arboles de decisión a partir de un conjunto de datos, utilizando una fórmula de predicción. Seguido utilizaremos Rpart.plot, este paquete nos sirve para poder graficar los resultados que nos ofrezca el rpart. Y por último la librería llamada ggplot2 la cual es un sistema de visualización de datos.

La manera de llamar estas librerías es descargando sus paquetes con la función (install.packages) y llamando las librerías después con la función (library) como se muestra a continuación:

```

install.packages("caret")
install.packages("ggplot2")
install.packages("rpart")
install.packages("rpart.plot")
install.packages("randomforest")
library(caret)
library(ggplot2)
library(rpart)
library(rpart.plot)

```

Una vez tengamos la base de datos lista y los paquetes y librerías necesarios el siguiente paso será cambiar los valores por valores numéricos, debemos convertir la variable a factor, a continuación, haremos uso de la función (as.numeric) para que las variables sean vistas como variables numéricas. Para realizar el árbol de decisión para la clasificación es importante que la variable sea categórica.

```

Suscriptores$Valor_Neto <- as.numeric(Suscriptores$Valor_Neto)
Suscriptores$Estado <- factor(Suscriptores$Estado)

```

Una vez se encuentren en numéricas vamos a plantar la semilla, lo cual es el punto de partida en el desarrollo de procesos alineados a árboles, se pueden escoger muchos tipos de semillas, pero para este caso que se trata de un proceso para clasificación la mejor semilla es la (123), y para plantarla se usa la función:

Set.seed(...) en donde los puntos suspensivos son cambiados por la semilla a usar.

En R, plantar una semilla significa iniciar un generador de números pseudoaleatorios, ya que la mayoría de los métodos de simulación requiere un generador de estos números. Para conseguir las secuencias de los números anteriormente mencionados se requiere un algoritmo recursivo llamado generador de números aleatorios.

$$x_i = x_{i-1}, x_{i-2}, x_{i-3}, \dots, x_{i-k}$$

Donde K es el orden del generador y $(x_0, x_1, x_2, \dots, x_{k-1})$ es la semilla (o el estado inicial del generador).

Lo siguiente será entrenar los datos que se posean, por lo que se necesita la siguiente línea de código:

```

X = createDataPartition(Y, p=0.#, list = FALSE)

```

En donde la letra X será cambiada por un nombre a nuestra elección, ya que por este se reconocerá el entrenamiento de nuestros datos. La letra Y será cambiada por los datos que vamos a utilizar, luego tenemos $p=0.\#$, el cual sirve para realizar la partición de los datos los cuales serán utilizados para el entrenamiento

El siguiente paso será crear el bosque aleatorio, para lo que usaremos la línea de código:

```
Nombre.a.asignar =  
randomForest(x=Base.de.datos[nombre.del.entrenamiento,1:4],  
y=base.de.datos[nombre.del.entrenamiento,5], ntree = #000, keep.forest = TRUE)
```

Donde primero asignaremos un nombre para nuestro bosque, se asignarán la cantidad de árboles con $ntree=###$ y se terminará la línea del código con $keep.forest=TRUE$.

Por consiguiente, se realizarán las pruebas con el siguiente código:

```
Nombre.a.asignar = predict (nombre.de.bosque, base.de.datos[  
nombre.del.entrenamiento,])
```

En donde se asignará un nombre para la predicción, finalmente se prosigue por el `predict`, que consiste en el nombre del bosque formado anteriormente, el nombre de la base de datos junto con las pruebas del entrenamiento, a las cuales accedemos con el símbolo “-“ y posterior el nombre de nuestro entrenamiento.

Resultados

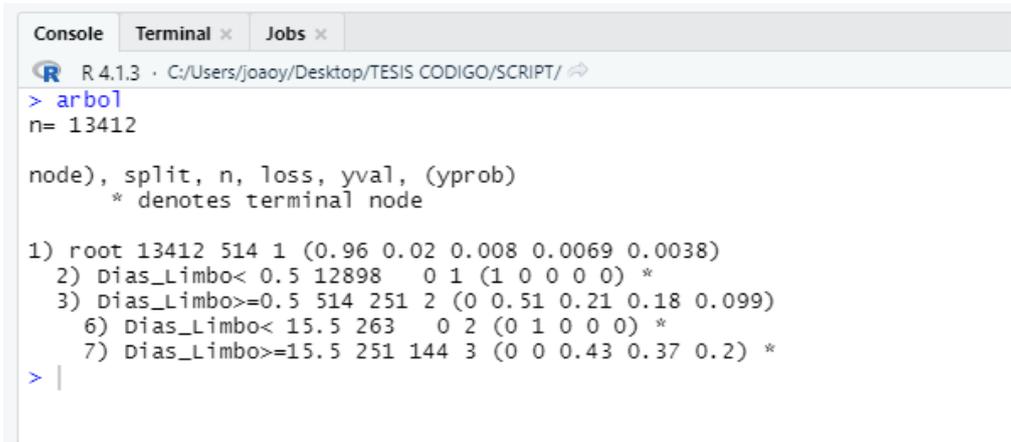
Mediante los modelos usados en esta investigación, se logró realizar un estudio a fondo de la fidelidad de los suscriptores del universo. Para empezar, se inserta una captura del árbol de decisión hecho al inicio, el cuál es el resultado de la siguiente línea de código:

```
arbol= rpart(Estado~., data = Suscriptores[Entrenamiento,],  
  
            method = "class", control = rpart.control(minsplit = 300, cp=0.01))
```

Figura

9.

Árbol de Decisión Generado en RStudio.



```
Console Terminal x Jobs x  
R 4.1.3 · C:/Users/joaoy/Desktop/TESIS CODIGO/SCRIPT/ ↗  
> arbol  
n= 13412  
  
node), split, n, loss, yval, (yprob)  
* denotes terminal node  
  
1) root 13412 514 1 (0.96 0.02 0.008 0.0069 0.0038)  
2) Dias_Limbo< 0.5 12898 0 1 (1 0 0 0 0) *  
3) Dias_Limbo>=0.5 514 251 2 (0 0.51 0.21 0.18 0.099)  
6) Dias_Limbo< 15.5 263 0 2 (0 1 0 0 0) *  
7) Dias_Limbo>=15.5 251 144 3 (0 0 0.43 0.37 0.2) *  
> |
```

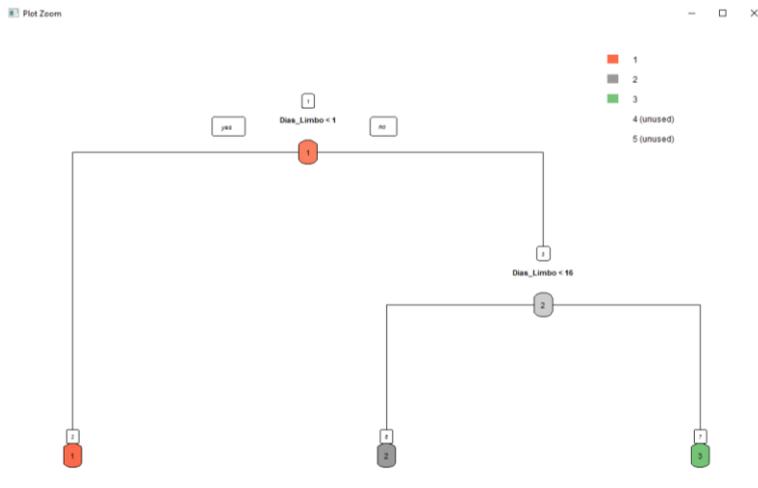
Al usar `rpart.plot`, es posible ver el árbol generado por RStudio de forma gráfica, de esa forma es más sencillo y más amigable a la vista.

```
rpart.plot(arbol, type = 1, digits = -1, extra = 0, cex = 0.7, nn= TRUE,  
  
          fallen.leaves = TRUE )
```

Figura

10.

Forma Gráfica del Árbol Generado por RStudio



Nuestra matriz de confusión nos arroja un resultado satisfactorio a simple vista, todos los suscriptores están correctamente clasificados según los criterios que se impusieron en la columna de “Estado”, donde: el 1 eran suscriptores actualmente activos, el 2 representa una alta probabilidad de que los suscriptores renueven suscripción, 3 una media probabilidad de que eso suceda, 4 baja probabilidad de que renueven suscripción y 5 nula probabilidad de que eso suceda.

Figura

11.

Matriz de Confusión.

```
> Matriz
      Predicho
Actual  1    2    3    4    5
  1 12898    0    0    0    0
  2     0   263    0    0    0
  3     0    0   107    0    0
  4     0    0    0    93    0
  5     0    0    0    0   51
```

Por lo que los resultados arrojan una correcta clasificación de los suscriptores en cuanto al estado que se encuentran.

También se optó por realizar y evaluar la probabilidad de que los suscriptores cambien de de estado, es decir, se evaluó la probabilidad de que estando inactivos, ya

sea en estado 2, 3, 4 o 5, estos pasen a un estado activo o en su defecto un estado distinto del que se encuentran.

Análisis de Resultados

Los datos al ser recopilados en un archivo de Excel, se encuentra guardado como formato .csv ya que este formato es el que permite una estructura tabular al momento de proyectar tablas y valores en RStudio. La base tiene como nombre “Suscriptores” en honor a la investigación a realizar.

Mediante los resultados presentados se evidenció que los suscriptores del universo, cuentan con una fidelidad extremadamente alta lo que resultó gracias a la evaluación del estado de los mismos. Las probabilidades de cambio de estado, demostraron que, será muy difícil que las personas suscritas cambien de un estado poco probable a renovar a una alta probabilidad de renovar suscripción, sin embargo, los suscriptores que realicen una suspensión temporal no de muchos días, tendrán una alta probabilidad de pasar de un estado inactivo a un estado activo, cabe mencionar que mientras pasan los días en el que el suscriptor está en el limbo, incrementarán las probabilidades de que este no retome la suscripción deseada.

En los anexos se podrá visualizar la base de datos con los resultados de las probabilidades y la predicción del estado.

Conclusiones

- Actualmente modelos de machine learning son de suma importancia a nivel empresarial si uno trata de destacar con el cliente o tener una ventaja competitiva con respecto a la competencia, existe una gran variedad de beneficios que nos pueden interesar como empresa y estos están basados en modelos de autoaprendizaje que nos proporcionan una visión más informada de la situación del sector de la empresa en el cual nos interesa llevar a cabo un proyecto de Inteligencia de Negocios. Entre los beneficios que puede brindar a las empresas es el ayudar a procesar grandes cantidades de datos, optimizar procesos, automatizar tareas, mejorar la eficiencia en general, tomar decisiones con mayor asertividad en base a información estudiada, descubrir patrones y tendencias en los datos, también puede ayudar a las empresas a personalizar su oferta y mejorar la experiencia del cliente. Sin embargo, el grado de importancia puede variar dependiendo del sector y las necesidades específicas de la empresa.
- El modelo que implementamos nos permitió tener una visión más real sobre la conducta de los clientes, mediante árbol de decisión y bosques aleatorios pudimos determinar que es necesario ejecutar estrategias independientes para cada grupo de clientes, ubicándolos de acuerdo con el tiempo que lleva desde que toma la decisión de suspender su suscripción, a través de dichas estrategias El Universo podrá tener una mejor retención de clientes, evitando bajas en el índice de suscriptores activos.

Recomendaciones

El Machine Learning se está convirtiendo cada vez más en una herramienta esencial en el ámbito empresarial de hoy en día, sin embargo es importante seguir algunas recomendaciones al utilizar estos procesos en una empresa, entre estas el tener bien planteada la problemática de la empresa es fundamental, esto incluye identificar el objetivo de la empresa y los datos que se disponen para realizar el modelo que se escoja, una vez se identifique la problemática otra parte importante es la preparación de los datos, limpiarlos y transformarlos para que estén en el formato que el modelo requiera.

Otro punto importante a la hora de desarrollar machine learning es seleccionar el modelo que se adecue a la problemática a la que nos vamos a enfrentar, y por último una vez que el modelo esté listo y nos arroje nuestros resultados también se recomienda realizarle pruebas al modelo después de emplearlo durante un tiempo determinado para corroborar que sigue trabajando bajo la misma eficiencia que nos brindó al comienzo del proyecto, tomando estas recomendaciones con suma precaución tendremos un mejor desarrollo y eficiencia del modelo y por ende un mejor resultado a la hora de implementar estrategias en la empresa.

Bibliografía

- Albornoz, M. B. (2020). Empatía. Retrieved from Empatía: <https://www.empatia.la/blogpost-ecuador-ia/>
- Alfonso, Y. (2019, 06 05). Gestipolis. Retrieved from Gestipolis: <https://www.gestipolis.com/servicio-que-es-caracteristicas-clasificacion-y-tipos/>
- Álvarez, I. I. (2022, 04 19). Computer World. Retrieved from Computer World: <https://www.computerworld.es/tendencias/la-inversion-en-inteligencia-artificial-en-estados-unidos-alcanzara-los-120000-millones-de-dolares-en-2025>
- Anonimo. (2021, 12 13). Mobile World Capital. Retrieved from Mobile World Capital: <https://mobileworldcapital.com/machine-learning-evolucion/>
- Arias, A. S. (2015, 09 04). Economipedia. Retrieved from Economipedia: <https://economipedia.com/definiciones/rentabilidad.html>
- Ariza, F., Rodríguez, J., & Fernández, A. (2018). Control estricto de matrices de confusión por medio de distribuciones multinomiales. *Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (21), 215-226. doi:<http://dx.doi.org/10.21138/GF.591>
- Betancourt, G. (2005). Las máquinas de soporte vectorial (SVMs). (27), 67-72.
- Bielschowsky, R. (2019, 09 03). Naciones Unidas. Retrieved from Naciones Unidas: <https://www.cepal.org/es/notas/cumbre-inteligencia-artificial-america-latina>
- Bonaccorso, G. (2017). *Machine learning algorithms*. Packt Publishing Ltd.
- Bonta, P., & Farber, M. (1994). 199 Preguntas sobre marketing y publicidad. In P. Bonta, & M. Farber, 199 Preguntas sobre marketing y publicidad (p. 178). Barcelo: Norma.
- Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi-Supervised Learning*. London: MIT Press Cambridge.
- Damodaran, A. (2007). *Probabilistic approaches: scenario analysis, decision trees and simulations*. New York: Stern School of Business.
- Doldán, F. (2007). Redes bayesianas y riesgo operacional. *Revista Galega de Economía*, 16, 1-16.

- Echavarría, M., Urrego, L., & Rodas, C. (2013). Aplicación de Árboles de Decisión para la valoración de compañías de minería aluvial. Medellín: Universidad EAFIT.
- Espinosa-Zúñiga, J. J. (2020, 12 02). Scielo. Retrieved from Scielo: https://www.scielo.org.mx/scielo.php?pid=S1405-77432020000300002&script=sci_arttext
- Garay, C. C. (19, 05 08). Logicalis. Retrieved from Logicalis: <https://blog.es.logicalis.com/analytics/presente-del-machine-learning-en-europa>
- Gironés, J. (2016). Algoritmos. España: Universidad Oberta de Cataluña.
- Godoy, Á. (2017). Técnicas de aprendizaje de máquina utilizadas para la minería de texto. 31(71), 103-126.
- Goyzueta, S. (2015). Big Data Marketing: una aproximación. Revista Perspectivas(35), 147.158.
- Gurney, K. (1997). An introduction to neural networks. London: Routledge.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. Electronic Markets(31), 685-695.
- López, J. (2009). Algoritmos y Programación. Colombia.
- López, J. F. (2021, 08). UISEK. Retrieved from UISEK: <https://repositorio.uisek.edu.ec/bitstream/123456789/4256/1/Freire-Lopez-Juan.pdf>
- Lu, F., & Bai, Q. (2010). A Refined Weighted K-Nearest Neighbors Algorithm for Text Categorization. IEEE, 326-330.
- Madhulatha, S. (2012). An overview on clustering methods. IOSR Journal of Engineering, 2(4), 719-725.
- Matich, D. (2001). Redes neuronales: conceptos básicos y aplicaciones. Universidad Tecnológica Nacional.
- Micó, J. L. (2018, 08 04). La Vanguardia. Retrieved from La Vanguardia: <https://www.lavanguardia.com/tecnologia/20180811/451181451132/inteligencia-artificial-estados-unidos-china.html>
- Mitchell, T. (1997). Machine Learning. McGraw-Hill.

- Morales, F. C. (2021, 02 14). Economipedia. Retrieved from Economipedia: <https://economipedia.com/definiciones/punto-de-venta.html>
- Moret, B. M. (1982). Decision Trees and Diagrams. ACM.
- Morisette, L., & Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Maths. 9(1), 15-24.
- Mueller, J. P., & Massaron, L. (2016). Data Science mit Python für Dummies. John Wiley & Sons.
- Nilsson, N. J. (1998). Introduction to Machine Learning. Stanford: Stanford University.
- Peiro, R. (2019, 12 09). Economipedia. Retrieved from Economipedia: <https://economipedia.com/definiciones/suscripcion.html>
- Recabal, P. (2009). Aprendizaje Reforzado orientado a la toma de decisiones en el fútbol robótico. Santiago de Chile: Universidad de Chile.
- Rivera, M. (2011). El papel de las redes bayesianas en la toma de decisiones. Bogotá: Universidad del Rosario.
- Rojas, E. M. (2020). Machine Learning: análisis de lenguajes de programación y herramientas para desarrollo. Revista Ibérica de Sistemas e Tecnologías de Informação, (E28), 586-599.
- Rouhiainen, L. (2018). Inteligencia Artificial: 101 cosas que debes saber hoy sobre nuestro futuro. Barcelona: Alienta Editorial.
- Salvador, C., & Cárdenas, J. (2020). Big Data: Información de vigilancia tecnológica. Lima: CONCYTEC.
- Santos, P. R. (2021). Tipos de aprendizaje en Machine Learning: supervisado y no
- Schiaffino, S. (2018). Clustering. ISISTAN.
- Shrestha, Y., Krishna, V., & Krogh, G. (2021). Augmenting organizational decision-making with deep learning algorithms. Journal of Business (123), 588-603.
- Silva, D. d. (2020, 12 09). Zendesk. Retrieved from Zendesk: <https://www.zendesk.com.mx/blog/perfiles-de-cliente/#:~:text=El-perfil-de-cliente-es-habitos-y-frecuencia-de-compra>.
- Silva, D. d. (2020, 10 16). Zendesk. Retrieved from Zendesk: <https://www.zendesk.com.mx/blog/servicio-post-entrega/supervisado>.

- Suresh, S., Sundararajan, N., & Savitha, R. (2012). Supervised Learning with Complex-valued Neural Networks. 1-29.
- Sutton, R., & Barto, A. (2015). Reinforcement Learning: An Introduction. London: The MIT Press Cambridge.
- Thompson, I. (2005, 12). Promonegocios. Retrieved from Promonegocios: <https://www.promonegocios.net/mercadotecnia/promocion-definicion-concepto.html>
- Unesco. (2019, 02 14). Retrieved from Unesco: <https://es.unesco.org/courier/2019-2/inteligencia-artificial-umbrales-africa>
- Valcárcel, V. (2004). Data Mining y el descubrimiento del conocimiento. Revista de investigación de Industrial Data, 7(2), 83-86.
- Vilca, D. A., & Camargo Ramos, J. (2021, 03 26). Uisek. Retrieved from Uisek: https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/16390/Aguilar_vd.pdf?sequence=1&isAllowed=y
- Villani, V., Pini, F., Leali, F., & Secchi, C. (2018). Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. Mechatronics, 55, 248-266.
- Vishwana, A. (2021, 10 22). NVIDIA. Retrieved from NVIDIA: <https://la.blogs.nvidia.com/2021/10/22/revolucion-ia-en-africa/>
- Wanjun, Y., & Xiaoguang, S. (2010). Research on text categorization based on machine learning. Advanced Management Science, 2, 253-255.
- Zhao, Y. (2013). R and Data Mining: Examples and Case Studies. Australia: Elseiver.

Anexos

Anexo 1.

Script completo en RStudio

```
setwd("C:/Users/joaoy/Desktop/TESIS CODIGO/SCRIPT")
Suscriptores <- read.csv2("../DATA/Suscriptores 2.csv")

library(caret)
library(ggplot2)
library(rpart)
library(rpart.plot)
install.packages("randomForest")

Suscriptores$Valor_Neto <- as.numeric(Suscriptores$Valor_Neto)
Suscriptores$Estado <- factor(Suscriptores$Estado)

#para un arbol o cualquier derivado de un arbol de decision, se planta una semilla
set.seed(123)

#entrenamos los datos
Entrenamiento <- createDataPartition(Suscriptores$Estado, p=0.85, list = FALSE)

#creamos el arbol de clasificacion
arbol= rpart(Estado~., data = Suscriptores[Entrenamiento,],
            method = "class", control = rpart.control(minsplit = 300, cp=0.01))
arbol

#diagrama de arbol
rpart.plot(arbol, type = 1, digits = -1, extra = 0, cex = 0.7, nn= TRUE,
           fallen.leaves = TRUE )

library(randomForest)

#Armamos el bosque aleatorio
RandomTreeModel <- randomForest(x=Suscriptores[Entrenamiento, 1:9],
                                y=Suscriptores[Entrenamiento,10],
                                ntree = 3000, keep.forest = TRUE)

RandomTreeModel

#Paso 5.- Realizo la prediccion
Prediccion <- predict(RandomTreeModel,Suscriptores[Entrenamiento,])
Prediccion

#Paso 6.- Armo la matriz de confusion
Matriz <- table(Suscriptores[Entrenamiento,"Estado"], Prediccion, dnn = c("Actual", "Predicho"))
Matriz

#Probabilidad de éxito
Suscriptores[Entrenamiento,"Probabilidad"] <- predict(RandomTreeModel,Suscriptores[Entrenamiento,], type = "prob")

#Valores clasificados
Suscriptores[Entrenamiento,"Prediccion"] <- predict(RandomTreeModel,Suscriptores[Entrenamiento,])

table(Suscriptores$Prediccion)
summary(Suscriptores$Probabilidad)
```

Anexo 2.

Base de Datos con Probabilidad y Predicción.

Estado	Probabilidad.1	Probabilidad.2	Probabilidad.3	Probabilidad.4	Probabilidad.5	Predicción
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.20000000000	0.20000000000	0.20000000000	0.20000000000	0.20000000000	1
1	0.20000000000	0.20000000000	0.20000000000	0.20000000000	0.20000000000	1
1	0.20000000000	0.20000000000	0.20000000000	0.20000000000	0.20000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.9120000000	0.0856666667	0.0016666667	0.0003333333	0.0003333333	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
5	0.0010000000	0.0250000000	0.0230000000	0.1306666667	0.8203333333	5
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

4	0.0006666667	0.0526666667	0.0793333333	0.8293333333	0.0380000000	4
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.9996666667	0.0003333333	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
2	0.0003333333	0.9893333333	0.0026666667	0.0076666667	0.0000000000	2
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

5	0.0000000000	0.0193333333	0.0393333333	0.1300000000	0.8113333333	5
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.99966666667	0.00033333333	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
---	--------------	--------------	--------------	--------------	--------------	---

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.20000000000	0.20000000000	0.20000000000	0.20000000000	0.20000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
2	0.0073333333	0.9660000000	0.0250000000	0.0013333333	0.0003333333	2
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
3	0.0053333333	0.0483333333	0.9293333333	0.0153333333	0.0016666667	3
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
3	0.0013333333	0.0360000000	0.8890000000	0.0243333333	0.0493333333	3
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.20000000000	0.20000000000	0.20000000000	0.20000000000	0.20000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1
1	0.10000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.2000000000	0.2000000000	0.2000000000	0.2000000000	0.2000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1
1	0.1000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000	1

1	0.100000000000	0.000000000000	0.000000000000	0.000000000000	0.000000000000	1
1	0.100000000000	0.000000000000	0.000000000000	0.000000000000	0.000000000000	1
1	0.100000000000	0.000000000000	0.000000000000	0.000000000000	0.000000000000	1
1	0.100000000000	0.000000000000	0.000000000000	0.000000000000	0.000000000000	1
1	0.100000000000	0.000000000000	0.000000000000	0.000000000000	0.000000000000	1

DECLARACIÓN Y AUTORIZACIÓN

Nosotros, Astudillo Muñoz Juan Camilo con C.C: # 0926778713 y Villares Zambrano Jhon Michael, con C.C: # 0930551288 autores del trabajo de integración curricular: **Aplicación de Machine Learning para la clasificación, suscripción y abandono/retiro de la suscripción de la Compañía Anónima el Universo**, previo a la obtención del título de **Licenciado en Negocios Internacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaramos tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de integración curricular para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizamos a la SENESCYT a tener una copia del referido trabajo de integración curricular, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 7 de febrero del 2023

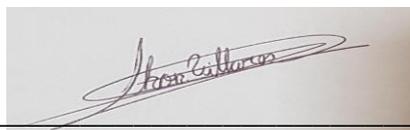
f. _____



Nombre: **Astudillo Muñoz, Juan Camilo**

C.C: **0926778713**

f. _____



Nombre: **Zambrano Villares, Jhon Michael**

C.C: **0930551288**

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TRABAJO DE INTEGRACIÓN CURRICULAR

TEMA Y SUBTEMA:	Aplicación de Machine Learning para la clasificación, suscripción y abandono/retiro de la suscripción de la Compañía Anónima el Universo.		
AUTOR(ES)	Juan Camilo, Astudillo Muñoz Jhon Michael, Villares Zambrano		
REVISOR(ES)/TUTOR(ES)	Carrera Buri, Félix Miguel, Mgs.		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Economía y Empresa		
CARRERA:	Negocios Internacionales		
TÍTULO OBTENIDO:	Licenciado en Negocios Internacionales		
FECHA DE PUBLICACIÓN:	07 de febrero del 2023	No. DE PÁGINAS:	87
ÁREAS TEMÁTICAS:	Economía, Innovaciones, Negocios		
PALABRAS CLAVES/ KEYWORDS:	Inteligencia de Negocios, Machine Learning, Árbol de Decisión, Bosque Aleatorio, Suscripción de Servicio.		
RESUMEN:	<p>La capacidad humana de crear cosas que incrementen la aptitud de los individuos crece constantemente, es por esto por lo que la tecnología da pasos gigantes y, corresponde a las personas, seguirle el paso y adaptarse, ya que cada año surgen nuevas herramientas con las que es posible explotar un poco más el talento que se posee. El machine learning, inteligencia de negocios, inteligencia artificial y distintos con las que se conoce esta herramienta, es una de estas herramientas que logra aumentar la seguridad al momento de realizar una decisión empresarial, dando una idea de lo que representaría el tomar dicha decisión, esto gracias a que da un resultado certero, disminuye el error al fallo y, acompañado de una buena implementación es una herramienta de enorme valor en las decisiones empresariales. Este proyecto tiene como finalidad demostrar que las herramientas como el machine learning pueden ayudar a prevenir situaciones que no se quieren en una empresa. Usando como ejemplo los suscriptores de la reconocida empresa C. A. El Universo. Para esto se usarán los modelos árbol de decisión, bosque aleatorio los cuales servirán para clasificar de forma correcta los suscriptores de esta empresa y evaluar la fidelidad de estos a dicha suscripción, según los resultados la empresa puede optar por una estrategia individual para cada persona y así, mantener al cliente suscrito al servicio, evitando una pérdida.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593 98 463 4250	E-mail Juan Camilo Astudillo Muñoz: c.a.13@outlook.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UIC):	Nombre: Román Bermeo, Cynthia Lizbeth		
	Teléfono: +593-984228698		
	E-mail: cynthia.roman@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			