**universidad tecnológica de aguascalientes**



**IMPLEMENTACIÓN DE TÉCNICAS DE INTELIGENCIA ARTIFICIAL EN LA CREACIÓN Y RECOMENDACIÓN DE RUTAS DE APRENDIZAJE**

**PRESENTA:**

**MELANIE ALEJANDRA VILLALOBOS ALFEREZ**

**AGUASCALIENTES, AGS AGOSTO 2020**

**PROYECTO REALIZADO EN THINCRS INC MÉXICO S. DE R.L. DE C.V..**

**ASESOR:**

**RICARDO ISRAEL ROQUE COVARRUBIAS**

**TUTOR:**

**CARLOS TOXTLI HERNANDEZ**

Agradecimientos o Dedicatorias

De las muchas personas que quiero agradecer es principalmente a Alberto Campos Hernández ya que él me recomendó en esta estadía, a pesar de mi poca experiencia, creyó en mí y en mis capacidades para poder aprender, también a Carlos mi tutor por aceptarme en su equipo de innovación, de tenerme mucha paciencia.  
  
Les agradezco por la oportunidad de poder darme cuenta sobre la como es trabajar en la vida real en base a la carrera, cabe resaltar que es mi primera vez estando en un equipo de trabajo, relacionado a la carrera a al que me quiero dedicar, el aprender un poco más de mi carrea, en poder ganar un poco más experiencia fue el mejor pago por todo esto y sin más solo me queda decir gracias al equipo, todo Thincrs por dejarme ser parte de su proyecto.

Tabla de Contenido

Contenido

[Introducción 1](#_Toc47228507)

[Capítulo I Generalidades de la empresa 2](#_Toc47228508)

[1. Datos generales 2](#_Toc47228509)

[2. Antecedentes Históricos 2](#_Toc47228510)

[3. Misión 2](#_Toc47228511)

[4. Visión 2](#_Toc47228512)

[5. Valores 2](#_Toc47228513)

[6. Productos o servicios que ofrece 3](#_Toc47228514)

[7. Organigrama 3](#_Toc47228515)

[Descripción del departamento donde se realizó la estadía 3](#_Toc47228516)

[Capítulo II: Planteamiento de la Estadía Profesional 4](#_Toc47228517)

[Descripción de la Problemática: 4](#_Toc47228518)

[Objetivos 4](#_Toc47228519)

[Justificación 4](#_Toc47228520)

[Resultados esperados 5](#_Toc47228521)

[Metodología a aplicar 5](#_Toc47228522)

[Estudio de viabilidad 6](#_Toc47228523)

[Estudio de factibilidad técnico 6](#_Toc47228524)

[Estudio de factibilidad operativa 7](#_Toc47228525)

[Estudio de factibilidad económica 9](#_Toc47228526)

[Resultado del estudio de viabilidad 10](#_Toc47228527)

[Capítulo III. Desarrollo de la estadía profesional. 11](#_Toc47228528)

[Diagrama entidad relación 11](#_Toc47228529)

[Machine learning 12](#_Toc47228530)

[Aprendizaje Supervisado 13](#_Toc47228531)

[Aprendizaje no supervisado 14](#_Toc47228532)

[Colectar Datos 16](#_Toc47228533)

[Preparar los datos 17](#_Toc47228534)

[Elegir el modelo de aprendizaje 19](#_Toc47228535)

[Perceptrón multicapa 21](#_Toc47228536)

[Entrenar máquina 22](#_Toc47228537)

[Proceso de entrenamiento 25](#_Toc47228538)

[Evaluaciones 26](#_Toc47228539)

[Parameter Tuning 29](#_Toc47228540)

[Predicción o inferencia 29](#_Toc47228541)

[Capítulo IV. Resultados y conclusiones 29](#_Toc47228542)

[Descripción de los resultados obtenidos conforme a los objetivos planteados en el proyecto 29](#_Toc47228543)

[Conclusiones. 30](#_Toc47228544)

[Recomendaciones. 31](#_Toc47228545)

[Anexos. 31](#_Toc47228546)

[Referencias Bibliográficas 34](#_Toc47228547)

# Introducción

Con el pasar del tiempo han emergido tecnologías que nos han ayudado a hacer las tareas o actividades de una manera más eficiente con la finalidad de poder mejorar procesos mediante recomendaciones automáticas que se usan en el área de aprendizaje es aquí donde la ciencia de machine learning ayuda a desarrollar técnicas para que las máquinas o aplicaciones que puedan aprender de los usuarios. Para lograr esto se debe tener una metodología para poder aplicar la ciencia de machine learning, seguir un modelo de aprendizaje que nos ayudara a poder evaluar a nuestra maquina y obtener los resultados mas cercanos mas reales en cuanto a sugerencias dependiendo de cada usuario.

Se quiere garantizar una buena experiencia para el usuario, navegando en un sistema intuitivo que permite fácilmente realizar evaluaciones y estas serán dirigidas mediante rutas de aprendizaje en las áreas de oportunidad encontradas para poder desarrollar las habilidades técnicas que se quieran reforzar o aprender, se validara el modelo implementado para el buen funcionamiento del sistema de recomendaciones y así poder dar una guía mas exacta al momento de dar una ruta de aprendizaje, validar que nuestro motor aprende a partir de información previa y a partir de ella poder dar una predicción mas concreta a la realidad.

En el presente documento se mostrará el proyecto de estadía en la empresa Thincrs, en el primer capítulo se dará de manera explícita la información general de la empresa, en el segundo capítulo se hablará sobre la problemática que se quiere resolver con dicho proyecto, en el tercer y cuarto capítulo se basará sobre las investigaciones y actividades asignadas en el transcurso de la estadía profesional y se dará el resultado final del proyecto.

# Capítulo I Generalidades de la empresa

## Datos generales

Thincrs fue fundado en 2016 por empresarios de IT con más de 15 años de experiencia, con el objetivo de reducir la brecha entre las tendencias y los talentos. Hasta la fecha, han preparado y posicionado a más de 400 talentos en México.

## Antecedentes Históricos

**Thincrs y Udemy quieren digitalizar la educación de Jalisco**

Thincrs, un startup mexicano que se encarga de evaluar, desarrollar y validar habilidades digitales de estudiantes, se alió con la plataforma Udemy y el Gobierno de Jalisco para acercar a miles de estudiantes el acceso a cursos en línea para fortalecer sus habilidades digitales.

La educación en línea se ha vuelto esencial en el mundo y en México, y más, en época de cuarentena. En este contexto tanto instituciones privadas como de gobierno han comenzado a responder con soluciones para mantener a los alumnos con sus actividades académicas al corriente.

## Misión

Ayudar a las personas a adquirir el conocimiento y habilidades que las industrias exigen.

## Visión

Ser la empresa líder en capacitación de vanguardia y creación de talentos tecnológicos.

## Valores

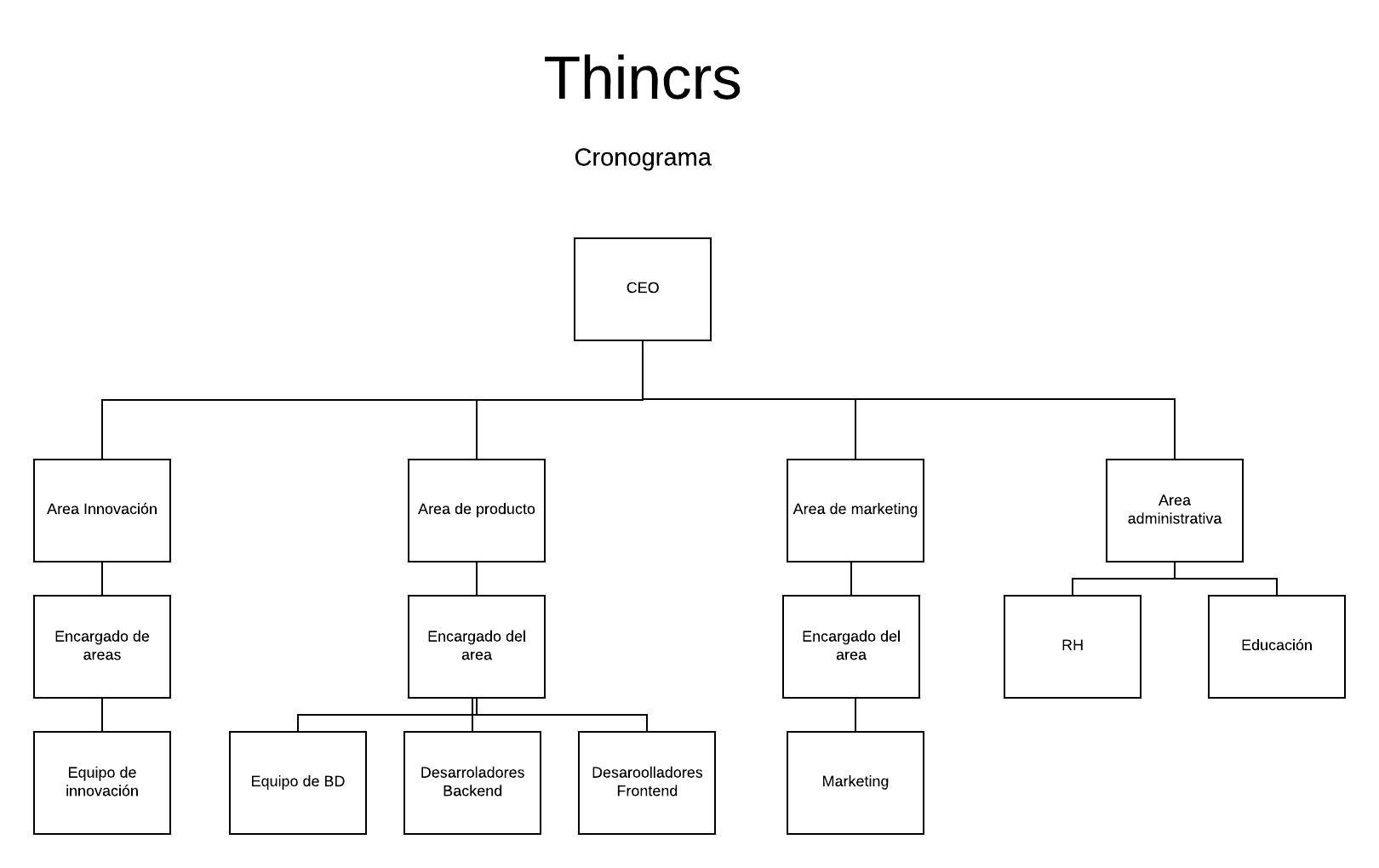
* Respeto
* Ética
* Trabajo en equipo
* Espíritu de servicio
* Responsabilidad

## Productos o servicios que ofrece

Una plataforma, adaptada a las necesidades específicas para el desarrollo del talento:

* Instituciones educativas ayuda a educar la fuerza laboral del futuro. Virtualiza la educación al desarrollar y validar las habilidades técnicas de estudiantes y maestros, en una única plataforma prescriptiva.
* Empresas que les permita las posibilidades de interrupción digital. Aumentando la gobernanza conectada en línea por los principales servidores públicos, maestros, ciudadanos y empresarios a la alfabetización técnica.

## Organigrama

****

## Descripción del departamento donde se realizó la estadía

Departamento de Innovación

En el área de innovación se trabaja con tecnologías para tener un sistema más proactivo.

En innovación se pasan procesos ya definidos por el área de producto que son los desarrolladores del sistema, por lo tanto en el departamento de innovación se encarga de analizar los comportamientos que se requieren de dichos procesos y poder controlarlo mediante un motor de IA (inteligencia artificial).   
En el departamento laboran:

* Carlos Toxtli Hernández
* Alberto Campos Hernández

# Capítulo II: Planteamiento de la Estadía Profesional

## Descripción de la Problemática:

Se quiere saber cómo interpretar y procesar las predicciones de una manera más eficiente.

Los métodos o modelos para implementar un sistema de machine learning llega a ser un poco exhaustiva, dichos modelos nos ayudan a poder tener una predicción lo más certera posible eso dependerá de las entradas y salidas de datos que se tengan en el proyecto.

## Objetivos

Validar el modelo de IA para garantizar la confiabilidad del sistema de aprendizaje automático.

## 

## Justificación

La razón por la que se debería tener un modelo válido de aprendizaje es para implementarlo según el funcionamiento para lo que se requiere, existen varios tipos de modelos, pero se debe tener uno en general para adecuarlo a los requerimientos del sistema.

## Resultados esperados

En el periodo de 4 meses se espera comprobar que la maquina inteligente tiene la precisión según el modelo elegido, ya entrenado con datos de prueba prestablecidos. Se quiere tener un porcentaje favorable de confianza en los resultados que nos arroje la maquina inteligente.

## Metodología a aplicar

Se describirán los 7 pasos genéricos para poder implementar un laboratorio de machine learning

1. Colectar Datos: Se hará una investigación y se obtendrá datos que se utilizarán para alimentar a la máquina, es importante importar la calidad de información que se consiga ya que esta impactará de manera directa a nuestro modelo de aprendizaje.
2. Prepara los datos: Se visualizará los datos obtenidos para hallar correlaciones entre cada uno de ellos, de la misma manera separa datos en dos grupos uno para entrenamiento y el otro para evaluación.
3. Elegir el modelo de aprendizaje: Se elige un modelo en base al objetivo que queremos cubrir, se utilizaran algoritmo de clasificación, predicción, regresión.
4. Entrenar a nuestra máquina: Se utilizará el set de datos de entrenamiento para ejecutar nuestra máquina, es importante inicializar los pesos de nuestro modelo aleatoriamente.
5. Evaluación: Se deberá comprobar que la máquina creada contra nuestro set de datos de evaluación que contiene entradas que el modelo desconoce y verificar la precisión de nuestro modelo ya entrenado.
6. Parameter Tuning: De no tener una buena evaluación en nuestras predicciones, se deberá retornar al paso de entrenamiento haciendo antes una nueva configuración de parámetros de nuestro modelo.
7. Predicción o inferencia: Se utilizara en el sector productivo al pasar los puntos de evaluación.

# Estudio de viabilidad

## Estudio de factibilidad técnico

En esta parte del documento se dará de manera detallada información sobre los recursos utilizados a lo largo del proyecto.

|  |  |  |
| --- | --- | --- |
| Equipo o producto | Cantidad | Descripción |
| Gestor de base de datos | 2 | Workbench 8.0 PHPMyAdmin |
| Sistemas operativos | 3 | Mac Ubuntu  Windows 10 |
| Laptops | 10 | Procesador AMD(8) o Intel (i5) para arriba, memoria ram DDR3, DDR4 512mb, mouse |
| Servidor sw | 2 | 4 procesadores virtuales, 16g ram, 30g de almacenamiento, sistema operativo .server 1604 |
| Servidor sw | 1 | 4g de ram, 2 procesadores virtuales |
| Hosting | 1 | 250 GB de almacenamiento, ancho de banda ilimitada |

## Estudio de factibilidad operativa

Permite a Thincrs poder hacer un análisis detallado de los usuarios que quieren desarrollar sus habilidades técnicas con su sistema de evaluaciones. Ofreciendo una ruta de retroalimentación dada por una IA.

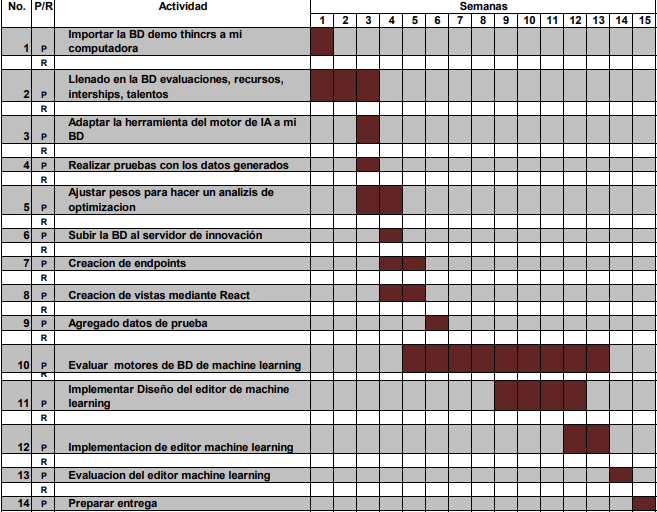
El sistema será intuitivo para el usuario con ayudas en los formularios para poder garantizar una buena experiencia al usuario.

La manera en que se implementan las capacitaciones es vía online con los encargados de departamento, estos te darán por parte del gobierno las herramientas para poder capacitarte en la técnica que se requiera implementar por medio de videos Udemy ellos te pasan una cuenta y tus haces los videos, posteriormente procedes a poner en practica lo que aprendiste y si tienes dudas solamente te comunicas con ellos por el Slack y te orientan.

Diariamente ellos están al día para cualquier duda que se llegue a presentar, por vía online, conferencias o chat.

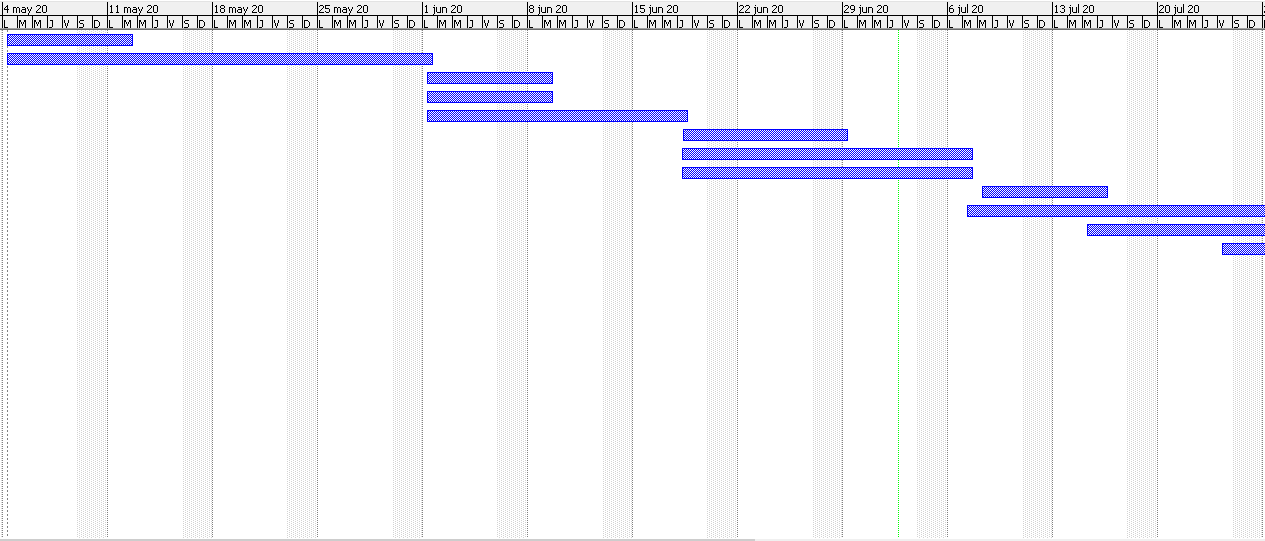
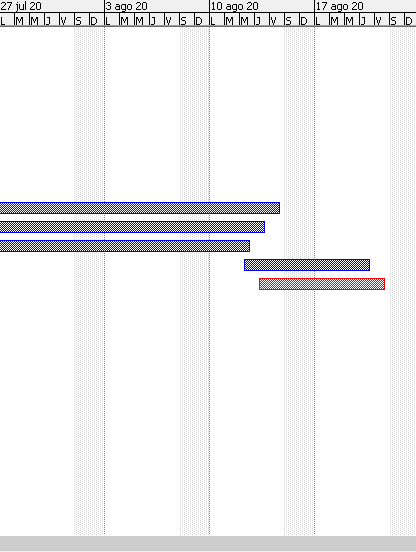
De la misma manera ellos te brindan las oportunidades para poder certificarte en alguna tecnología.

A continuación, mostraré las actividades que se desarrollarán a lo largo del proyecto:



En el siguiente diagrama de riesgos se identifica las actividades o procesos que están sujetos a un retraso en nuestro proyecto y poder cuantificar la probabilidad de que esto pasen en una la línea de tiempo según el periodo de vida del proyecto.

La línea roja representa las actividades que pondrían en riesgo a la entrega del proyecto final, se tomaran medidas de precaución para poder dejar tiempo extra en ese periodo para que no tenga impactos negativos en las entregas.



## 

## Estudio de factibilidad económica

Lista de costos

|  |  |  |  |
| --- | --- | --- | --- |
| Servicio o producto | Marca | Tipo | Costo al mes |
| Servidores | Amazon | T2 ½ | $384.8 |
| T2 | $7531.77 |
| Hosting | Godaddy | Dominio | $431.00 |

Servicios server se renuevan cada 6 meses

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Microsoft Azure | | AmazonWeb Services | |
| Mensual | Precio por hora | Precio mensual | Precio por hora | Precio mensual |
| 1 | $13.6 | $5032.0 | $5.22 | $1931.4 |
| 2 | $13.6 | $5032.0 | $5.22 | $1931.4 |
| 3 | $13.6 | $5032.0 | $5.22 | $1931.4 |
| 4 | $13.6 | $5032.0 | $5.22 | $1931.4 |
| 5 | $13.6 | $5032.0 | $5.22 | $1931.4 |
| 6 | $13.6 | $5032.0 | $5.22 | $1931.4 |

Usando AmazonWeb Services comparado contra Microsft Azure se tiene una diferencia en precios de las siguientes cifras:

|  |  |
| --- | --- |
| Diferencia por mes | 3, 100.6 |
| Diferencia por 6 meses | **18603.6** |

## 

## Resultado del estudio de viabilidad

Dentro del estudio de viabilidad se llegó a una conclusión de que las herramientas que se implementa a lo largo del proyecto son las más indicadas en costos, calidad, en costos porque se siente una capacidad elevada de almacenamiento por Amazon, este ofrece precios accesibles según a la capacidad y cada cierto tiempo se brindan cupones de descuento.

Los procesos que se llevan acabo para que el personal pueda trabajar y ser entrenado son las mas optimas, se tiene gente capacitada para poder asesorar las dudas al momento de algún problema en el desarrollo del sistema inteligente.

# 

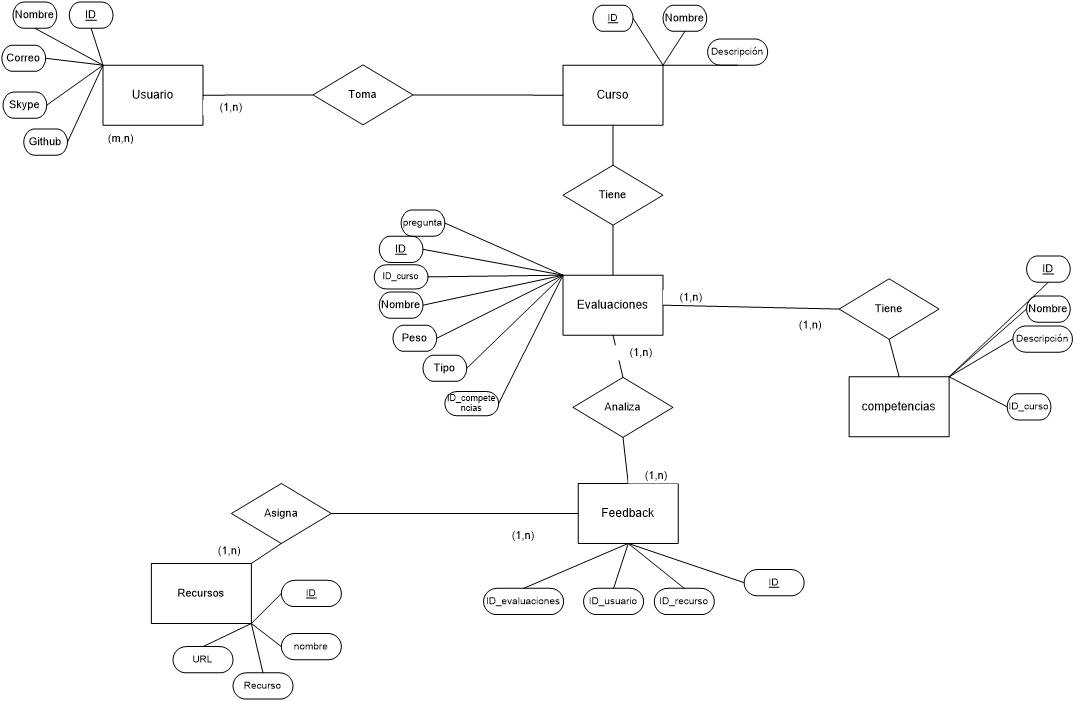
Por ahora los requerimientos del sistema se mantienen estables, lo que mas llega a costear es el hosting pero se compensa con el ahorro del pago de los servidores.

# Capítulo III. Desarrollo de la estadía profesional.

## Diagrama entidad relación

Descripción del sistema:

En Thincrs un usuario tiene un id único, nombre, correo, GitHub, Skype (opcional) este usuario toma un curso (basados en tecnologías como java, js, móvil, html, jQuery) que tiene evaluaciones, dichas evaluaciones son realizadas por los usuarios, las evaluaciones se basan en competencias que describen un objetivo cuyo propósito es resolver una problemática en relación a una tecnología.  
Las evaluaciones tienen una retroalimentación y en base a esas retroalimentaciones se dan recomendaciones de recursos a los usuarios.

****

## Machine learning

Esta tecnología que se ha innovado con el paso del tiempo se basa en el aprendizaje automático a partir de un algoritmo complejo donde se procesan cientos o miles de datos (Dataset) que se usarán para poder entrenar al sistema, estos datos históricos tienen como propiedad instancias de factores y características, con la finalidad de poder aprender de ellos y en base a un modelo de aprendizaje el cual se implementa con la finalidad de poder predecir eventos futuros.

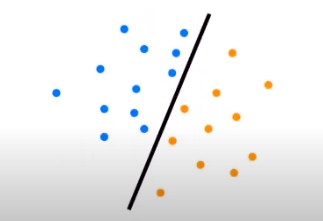
Los modelos son filtros de entradas o salidas donde las salidas conlleva la manera de clasificar a un dato según los patrones que se dé detectan en el entrenamiento.

Para entender un poco mas el concepto de modelos les hablare de algunos de ellos, pero primero cabe resaltar que la maquina inteligente tiene distintas formas de aprender, solo dos, cada una de ellas tiene un modelo distinto y distintas funciones, a continuación, les hablare de algunos de ellos para poder diferenciar sus características como modelos de aprendizaje:

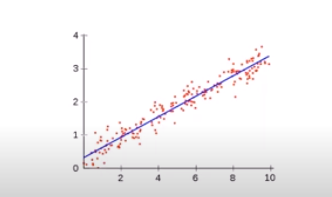
### Aprendizaje Supervisado

Dentro del aprendizaje supervisado se usa un algoritmo que aprende por medio de etiquetas las cuales se guarda información para después hacer previsiones futuras existen dos tipos de aprendizaje supervisado los cuales son:

* Clasificado: El algoritmo encuentra diferentes patrones y clasifica los elementos en diferentes grupos. Este modelo busca sacar conclusiones de los valores observados, ya que una o más entradas intentan predecir el valor de uno o más resultados. Un claro ejemplo para nuestro caso seria la clasificación de usuarios, cursos, evaluaciones y poder dividir dicha información por secciones o grupos.



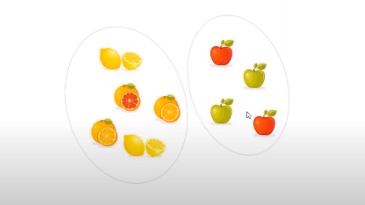
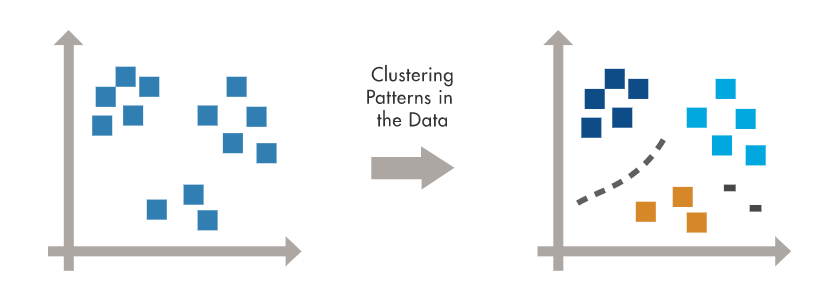
* Regresión: Tiene el objetivo de predecir valores continuos. Con el valor numérico de las etiquetas, se utilizan diferentes variables para obtener los datos que nos interesen. Este tipo de aprendizaje sirve para una serie de finalidades concretas como predecir el próximo curso que un usuario tomara.



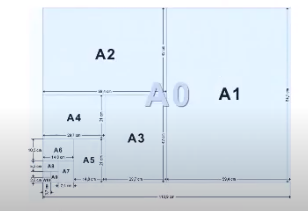
### Aprendizaje no supervisado

A diferencia del aprendizaje supervisado, en el no supervisado solo se le otorgan las características, ya que se trata de extraer conocimiento de conjuntos de datos, no existe una etiqueta. En este caso, el algoritmo cataloga los datos por similitud y crea grupos, sin tener la capacidad de definir cómo es cada individualidad de cada uno de los integrantes del grupo, existen dos modelos:

* Análisis de clúster: Esta extrae referencias de un grupo de datos sin etiquetas o sin clasificación. Se trata de clasificar por grupos homogéneos entre sí y los más heterogéneos para reducir la complejidad del sistema.

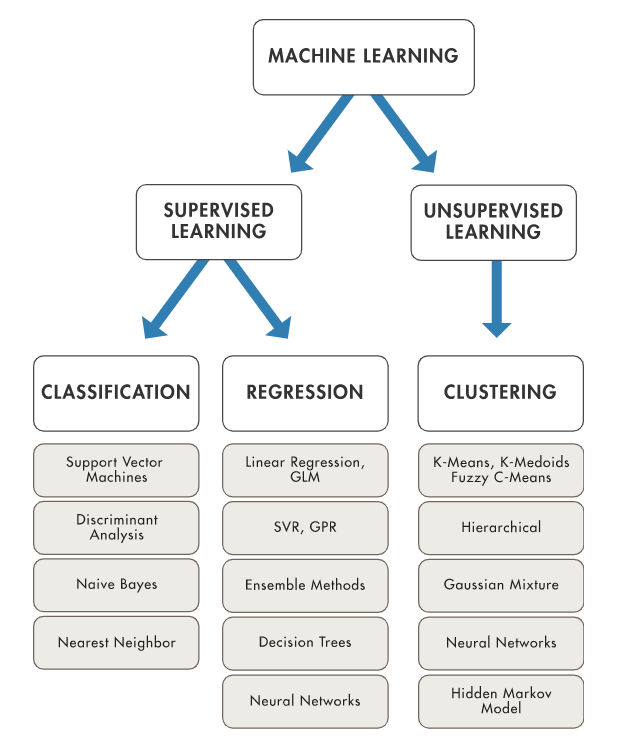
* Reducción de la dimensionalidad: Se trata de simplificar o reducir las variables aleatorias mediante la obtención de un conjunto de variables, así mismo eliminando variables irrelevantes.



Ambos métodos de aprendizaje se basan en una estructura para poder aprender y predecir que seria mediante un conjunto de datos filtrarlos por un vector de características y por un algoritmo de aprendizaje automático pasarlo a nuestro modelo predefinido y tener un resultado, solo que en el aprendizaje no supervisado no hay etiquetas para poder clasificar dichos datos, mas bien toma las características de estos datos y los agrupa por sus conjuntos de información o longitud, pesos etc.

La ciencia de machine learning tiene varias aplicaciones como en la medicina, en los bancos, en cualquier sitio web, aquí se aplicará en el área de educación con la finalidad de poder procesar una ruta de aprendizaje a los usuarios, la información previa que se recopila en sus evaluaciones y cursos tomados ayudarán a entrenar al sistema en base a cada usuario y poder predecir qué curso es más factible que este tome.

Dentro de cada método de aprendizaje hay mas funciones y nosotros debemos trabajar basado en uno, entonces ya mostrados los métodos y funciones de cada uno, de manera general se puede elegir uno valido para nuestro proyecto.



## Colectar Datos

Dada la cantidad masiva de datos que se necesitan para poder entrenar a nuestra máquina se utilizarán herramientas de apoyo con las que podremos implementar scripts, apis y llenar la BD.

Se plantea tener 100 usuarios los cuales harán una simulación de estudiantes de alto, medio y bajo nivel, cada usuario deberá ser tener una habilidad técnica en el ámbito de las tecnologías.

Lo que se usó fue

* Laravel: Con laravel se plantea el método de consumir una aplicación mediante PHP, se tienen más herramientas, funciones y librerías que nos pueden permitir poder hacer el llenado más dinámico de datos (factorys, migraciones), esto nos sirve para generar información aleatoria como nombres y correos.
* Cargas de excel: Las cargas de excel son hechas para información ya predefinida por la misma empresa como las evaluaciones, los cursos, las competencias. Y sobre un CVS poder cargar esa información por un proyecto de PHP.
* Scripts: Mediante los scripts podemos manejar el modelo de relación, que también se podría por laravel y el modelo de eloquentem, solo para tener una congruencia entre la información recolectada, se usarán scripts para las tablas pivote que suelen unir a más de una tabla.

Como apoyo para poder colectar nuestros datos se usa un análisis exploratorio de datos llamado EDA DataScience o análisis exploratorio de datos, descubierto por John W. Tukey que consiste en procesas la información por etapas, donde primero se recolectan los datos que vamos a necesitas, después por medio de un programa en este caso implementado por Laravel o un cualquier programa que colecte datos, hacer una limpieza de datos incensarios y procesarlos para poder llenar nuestra BD, lo que nos ayuda EDA seria en obtener las características importantes de los datos.

## Preparar los datos

Datos de aprendizaje

Los datos de aprendizaje serán aquellos que nuestra maquina inteligente tomara como referencia para poder aprender.

Estos datos deben ser coherentes para poder probarlos, todo es por ensayo y error, no se trata de llenar información al azar eso solamente complicaría las cosas al momento de probar.

A continuación, se mostrar que datos de aprendizaje tomará la maquina inteligente

Evaluaciones

Cada usuario tendrá una personalidad dentro de la BD estará será definida:

* Su habilidad técnica en una tecnología (BD, POO, Programación web)
* Su nivel de dominio sobre esa habilidad
* Su escolaridad
* Su localidad
* La edad

En el siguiente diagrama se muestra los tres cursos que engloban nuestras evaluaciones y como se caracteriza por competencias al menos por estas pruebas, solo se plantean 3 cursos.

Se plantea que cuando se lleve a producción habrá mas cursos para poder escoger.

Competencias para evaluar en base a una problemática de la vida real

Java

WEB

Movil

Los usuarios se clasificaron en tres niveles dependiendo a una tecnología dominada o no dominada, un usuario puede ser nivel básico, intermedio o avanzado a partir de sus evaluaciones se definirá el nivel. Serán creados 100 usuarios, existirán 34 usuarios de nivel básico, 33 intermedios y 33 avanzados.

Datos de evaluación modelo

Los siguientes datos son los cursos que son dados por el departamento de educación, un curso es un bloque de educación en relación a una tecnología.  
  
Las evaluaciones son el conjunto de categorías que llevan una serie de preguntas dada por las competencias y los cursos, dichas evaluaciones tienen un peso (weigth) que nos ayudará a poder clasificar la importancia de cada pregunta, hay desde las básicas hasta preguntas difíciles.

Las competencias son la razón o el objetivo de nuestra evaluación, conllevan un reto a poder lograr o lo que se pretende que el usuario deba hacer.  
  
El feedback se basa en los resultados de las evaluaciones, aquí se guardan las respuestas erróneas de cada usuario por evaluación y posteriormente tenemos los recursos donde se guardan en una tabla con cursos que pueden ayudar a reforzar a los usuarios en los temas que llegaron a tener mal en las evaluaciones, por eso es importante tener un control de los pesos que se les da a las evaluaciones y preguntas.

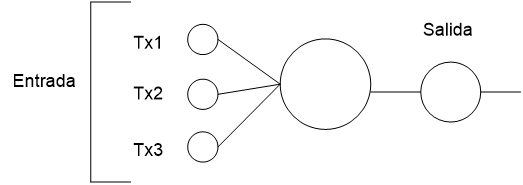
Una de las tablas secundarias del feedback serán como ayuda para tener más control en cómo van avanzando los usuarios con los recursos asignados a un usuario y cuanto progreso se tiene de dichos recursos, si es que los ve por completo o no los toma o si exitosamente este llega a completarlos, eso ayudara a la maquina inteligente para no saturar al usuario de recursos cuando este falle en su mayoría de evaluaciones.

La manera en que una evaluación se verá reflejada será por medio de las respuestas de los usuarios, esto significa que un usuario tiende a fallar en las preguntas y en base a eso las evaluaciones se irán haciendo más fáciles o difíciles.

## Elegir el modelo de aprendizaje

Red neuronal

Este modelo de aprendizaje supervisado se trabaja con vectores de entradas que devuelven un vector de salida, estos vectores están manejados por neuronas, dichas neuronas están conectadas a más neuronas, que llevan a una salida. La activación de ellas será secuencial a las que tengan en relación.



Antes de explicar cualquier otra cosa primero definiremos lo que es un perceptrón, no es más que una neurona artificial en este caso, es una unidad o calculo que intenta interpretar un comportamiento de una neurona natural, es un algoritmo que genera un criterio para hacer un grupo de datos.

A close up of a clock

Description automatically generated

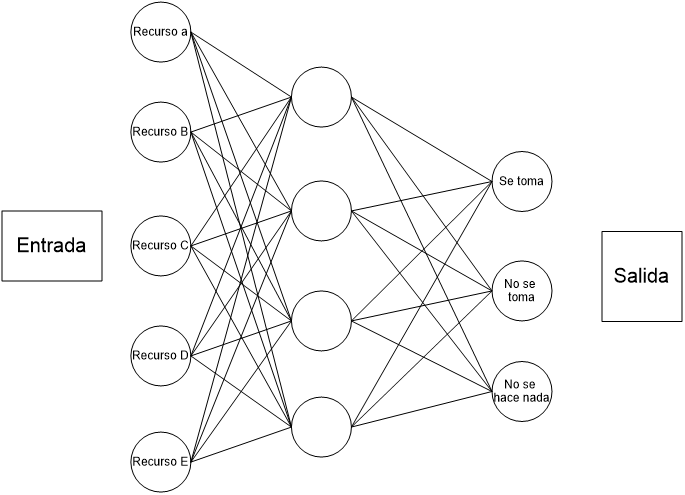
Hay distintas funciones de activación que se pueden usar al implementar un modelo neuronal, a lo largo del proyecto se implementará la función Umbral que se explicará en un momento:

* Función Umbral o Threshol

En las funciones umbrales reciben un valor dicho valor se procesa para dar una salida, la neurona de salida será la última neurona del lado derecho como se muestra a continuación.

A partir de nuestras salidas se pueden tomar decisiones de si tomo un curso o no tomo un curso que podría pasar, esto nos ayuda a clasificar nuestros datos previos.

La clasificación en este caso nos ayuda a entender los patrones de comportamiento de un usuario a lo largo del trayecto del sistema cuando este hace sus evaluaciones.



La manera en que se trabajará en este proyecto será mediante pasos por cada Sprint según nuestra metodología escogida, para poder evaluar el buen funcionamiento del sistema de IA existente necesitamos primero tener datos con los cuales nos ayudarán aprender y optimizar.

La razón por la que la red neuronal se implementa en este proyecto es porque existen distintos varios tipos de red neuronal y estas se pueden llegar acoplar cuando se trata de mucha información que filtrar, como entes caso.  
El tipo de rede neuronal que se verá es:

### Perceptrón multicapa

Como dije anterior mente las redes neuronales trabajan a partir de funciones existen las lineales y no líneas y estas se diferencias del conjunto de datos relacionados entre si, el perceptrón multicapa trabaja con funciones no lineales, podemos aproximar funciones no líneas a partir del perceptrón multicapa, a partir de esta podremos clasificar nuestros datos.

A picture containing game

Description automatically generated

La capa de entrada conecta con el exterior, cada neurona se corresponde con cada una de las variables de la red.

Las capas ocultas son el conjunto de capas en las que se activa una salida y esta procede a la suma ponderada de las activaciones de las capas anteriores conectadas, más sus correspondientes.

La capa de salida conecta a las capas ocultas con la salida de la red que proporciona los resultados esperados de una propagación que es cálculo de la salida de la red desde los valores de entrada. El aprendizaje empieza cuando los errores obtenidos a la salida del perceptrón se van propagando hacia atrás.

## Entrenar máquina

Bueno como mencionamos al inicio del documento, nuestra máquina aprende con el modelo de aprendizaje supervisado que consiste en la máquina inteligente colecte datos clasificados mediante algoritmos, clasificando los datos históricos tienen una serie de parámetros con los cuales se podrá entrenar nuestra máquina.

El proceso de aprendizaje se basa en tres funciones:

* Forward propagation: es el proceso por el cual una red neuronal envía su entrada a través de sus capas hacia la salida.
* Calcular el coste de la función: el proceso donde se calcula el peso de las neuronas.
* Backpropagation: es el proceso que busca ajustar los pesos de cada neurona de tal manera que se minimice el error.

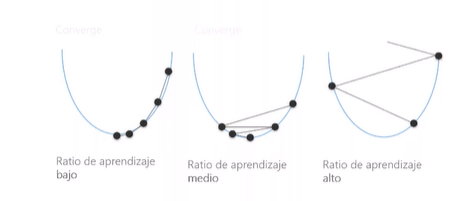
Se quiere aprender de un usuario el cual se le pondera de la siguiente manera

Entonces queremos la suma de las calificaciones de tales notas y mediante pruebas y errores podemos llegar al resultado correcto, mediante las funciones Forward propagation donde se ponen los pesos o calificaciones, se hace el cálculo, se reajustan los pesos mediante un descenso de gradiente hasta que se llega al resultado correcto.

8 8 8

66 6

El descenso del gradiente no es más que la ratio de aprendizaje que va entre nodos o secciones para poder reajustar los pesos de los nodos, estos van a prendiendo al final de cada nodo con el objetivo de modificar los pesos de las conexiones para que el valor estimado de la red se asemeje a algo más real, esto se define como función de gradriente del error.



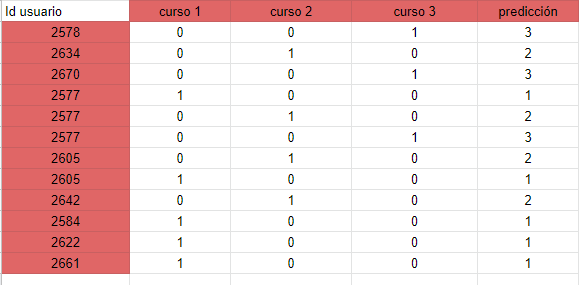
Por cada nodo que se quiera aprender se debe ajustar los pesos y se debe definir el tipo de ratio de aprendizaje para evitar desvíos infinitos, que eso nos llevaría a resultados no complacientes al momento de dar una predicción, los siguientes datos conllevan un beso a valorar

* Las preguntas: por cada pregunta se tiene un peso ya que se miden por dificultad entonces al momento de que un usuario x hace una evaluación esta toma el peso según a las preguntas.

Otro ejemplo donde nuestra maquina inteligente aprende es cuando los usuarios tomando un curso recomendado, la maquina solo interpreta 0 y 1.  
Pero el sistema es mas complejo aun ya que se maneja por matrices los datos almacenados por medio de consultas a la BD y este los procesa y filtra dando estos resultados como ejemplo:

Tenemos que un usuario toma el curso 3 y el segundo usuario toma el curso dos

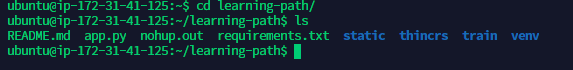
Entonces en base al historial de cursos tomados este aprende cuales son los que mas se repiten y se basa en eso para poder dar una predicción de cuál sería el siguiente curso que vería otro talento o este casi el usuario.



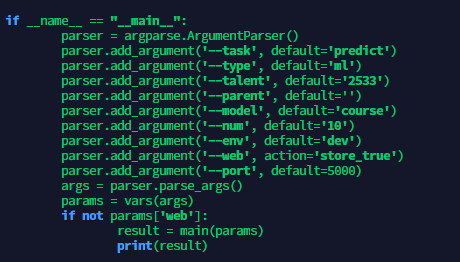
### Proceso de entrenamiento

Lo primero será clonar el proyecto desde gitlab a nuestra maquina por el comando git clone url de git.

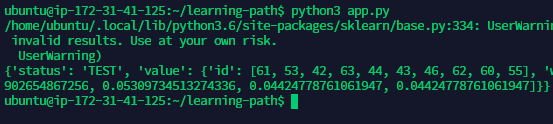
Posterior a eso tendremos que instalar las librerías dependientes desde un script hacer jalar el motor en nuestra maquina y configurarla se la siguiente manera en el archivo app.py



Dentro se puede configurar de manera manual, la base de datos local, el usuario del cual se quiere aprender, modelo donde definiremos si es por curso o evaluación, el tipo de motor si es de aprendizaje, optimización.



Para poder comprobar que nuestro motor funciona solo corremos la aplicación con el comando app.py



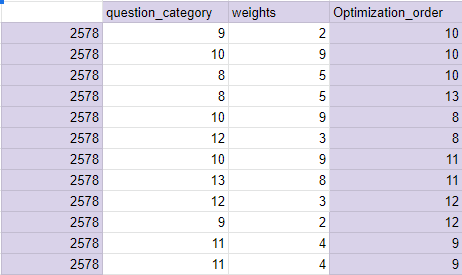
## Evaluaciones

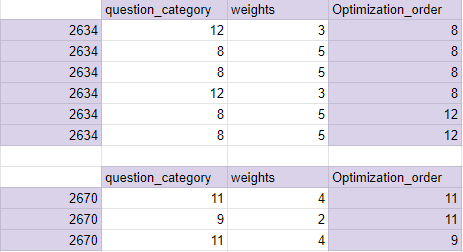
Para poder realizar las evaluaciones de nuestro maquina inteligente vamos tomar un 80% de usuarios para entrenar y el 20 % para evaluar cada Categoría es una evaluación.

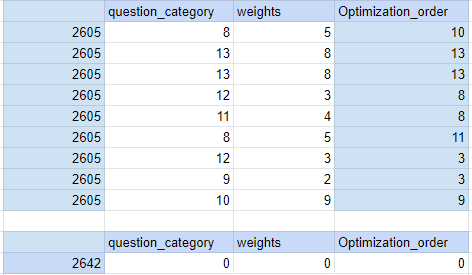
La herramienta donde se implementan las pruebas es en un servidor el cual yo me conecto por la herramienta Termius, que es un cliente SSH multiplataforma, que se ejecuta en computadoras de escritorio y dispositivos móviles. Este cliente SSH permite organizar hosts. El motor de inteligente esta elaborado en Python y por medio de librerías como Panda, Flask, skalearn ; es que este proyecto funciona.

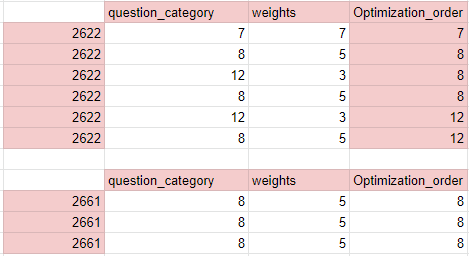
Se puede probar de dos maneras, desde una terminal o un api desde navegador, solo basta con correr el programa con App.py y se agreguen los parámetros solicitados.

En base a las evaluaciones, por cada usuario que fallo en una pregunta, se le manda un feedback dependiendo de la competencia que tenga relacionada, cada pregunta tiene un peso y se quiere la optimización de cada usuario por pregunta incorrecta de las preguntas con mayor valor a menos valor mis resultados, de este modo según las preguntas incorrectas se pueden redirigir en relación a los cursos para que el usuario tome, fueron los siguientes:

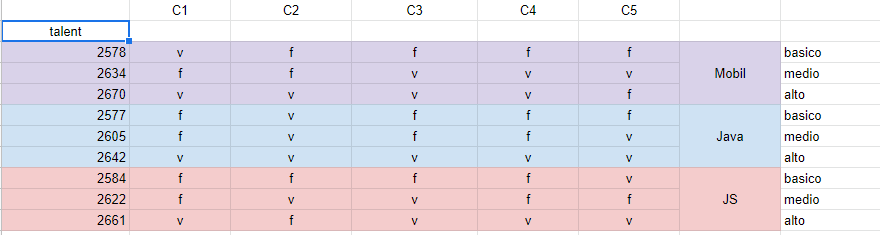




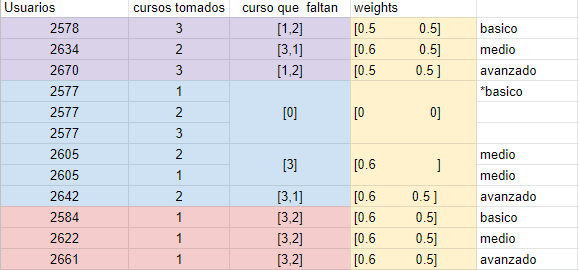




Aquí se muestran los usuarios en una tabla de verdad, donde se plasma como fueron evaluados.



Mis resultados son los siguientes: tengo casos por cada usuario según su habilidad un caso distinto para poder probar desde el estudiante más avanzado que logro tener evaluaciones perfectas, hasta el usuario que estuvo mal en todo.



Si alcanzamos un 90% o más podremos tener una buena confianza en los resultados que nos otorga el modelo.

## Parameter Tuning

En el caso de no tener una buena evaluación en nuestras predicciones, se deberá retornar al paso de entrenamiento haciendo antes una nueva configuración de parámetros de nuestro modelo.

Este “tuneo” sigue siendo más un arte que una ciencia y se ira mejorando a medida que experimentamos. Al finalizar esta etapa se vuele a dar una evaluación del modelo.

## Predicción o inferencia

Una vez pasado por el paso de evaluación de maneta exitosa, se transfiere esta maquina inteligente al sector productico con personas reales, estas pruebas se implementarán cuando el área de producto finalice la integración continua de este módulo.

# Capítulo IV. Resultados y conclusiones

Descripción de los resultados obtenidos conforme a los objetivos planteados en el proyecto

Dentro de las pruebas implementadas, los resultados obtenidos con los datos generados de manera coherente se pudieron llegar a un porcentaje favorable o de confianza para validar el modelo de aprendizaje y ajustar las matrices para poder clasificar nuestros datos de entrada.

El ajuste de la ratio es de medio aprendizaje es el adecuado para evitar desfases al momento de ajustar nuevos cálculos cuando la maquina se le entrene.   
Hubo un desfase de error del 5%

## Conclusiones.

Dentro de la ciencia de machine learning se plantea la automatización que esto significa a la conversión de uno o más procesos del sistema de Thincrs, y con la ayuda de esta ciencia se ahorra la capacidad de procesamiento de datos para el mejor consumo de los servicios.  
La capacidad de las maquinas inteligentes serian para la disminución de errores, ya que errar es de humano, pero configurando una máquina que pueda controlar procesos ya definidos es más certero tener una confianza en los resultados obtenidos, brinda un mejor servicio del sistema al cliente, concluyo que esta herramienta tiene el potencial para poder implementarse en cualquier sector laboral, desde el sector de aprendizaje hasta la medicina. Es una ciencia que esta evolucionando con el paso del tiempo y va de la mano con las tendencias de seguridad.

Los modelos que evalúan a las maquinas inteligentes son una base importante porque es la estructura y las funciones con las que vas a emplear tu proyecto, se tiene que acoplar a tus requerimientos, por ejemplo una red neuronal convolucional nos ayuda a clasificar imágenes por medio de patrones, en cambio los modelos como el diagrama de árbol, red neuronal pueden procesas adecuadamente este tipo de datos, por medio de sus pesos o características, cual sea el modelo elegido se tendrá mayor precisión al momento de hacer pruebas ya que cada una aprende de maneras distintas y algunas tienden a tener mayor porcentaje de error y eso podría ser muy contraproducente.

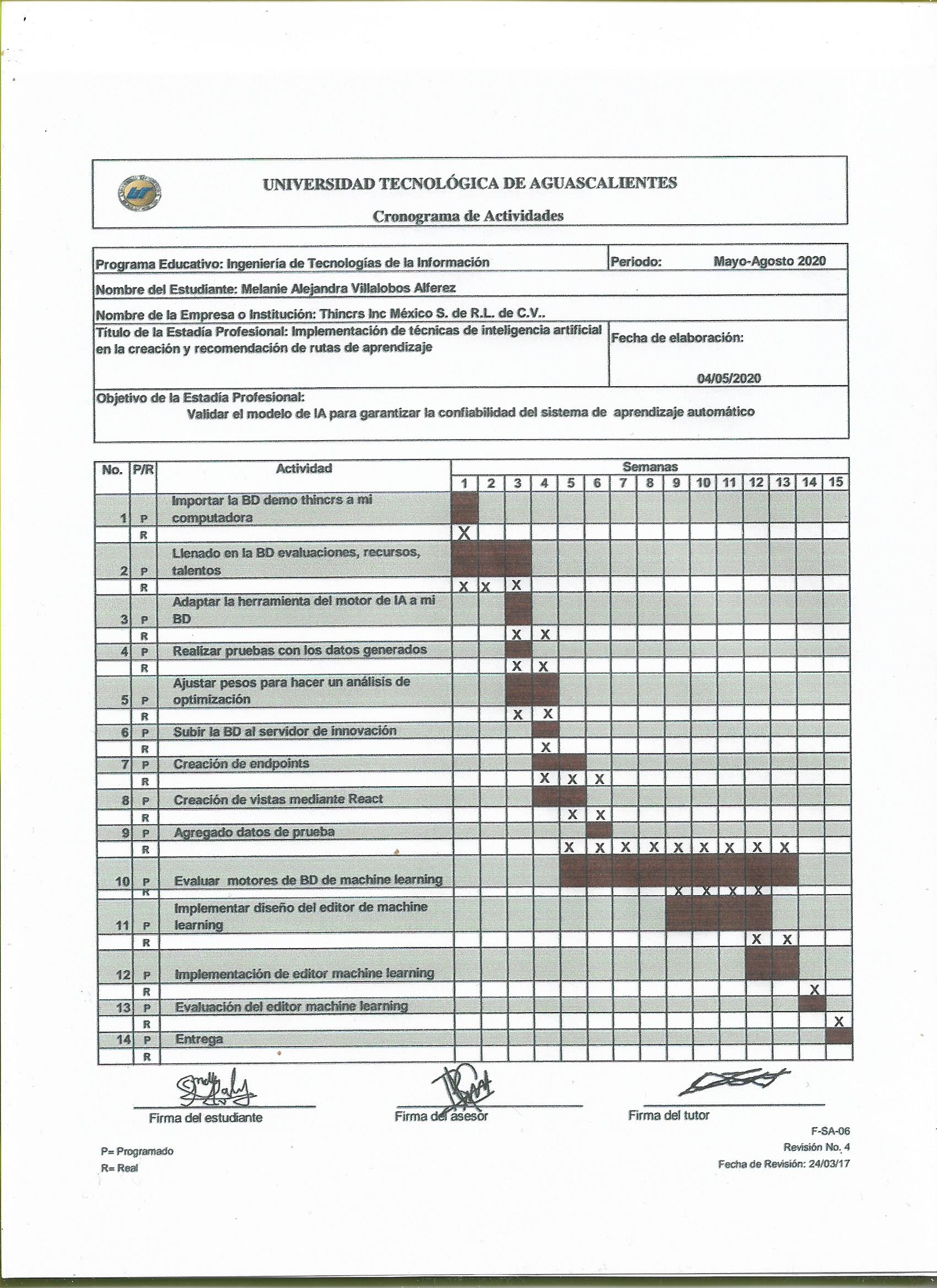
Al momento de colectar los datos si es muy recomendable que se use una herramienta de recolección de datos para no hacer tan laborioso el trabajo, ya hay muchos programas que hacen ese tipo de tareas, este tipo de herramientas funcionan cuando vas a hacer una colección de datos masivos, por ejemplo aquí pues no eran tanto datos como para usarla pero esta bien dejarlo en claro que es una herramienta de apoyo.

## Recomendaciones.

Bueno por parte mía para la empresa pues solamente tengo una recomendación muy general, que seria respetar un poco los tiempos en que se plantean las entregas ya que el querer entregar algo bien funcionando pues tiene su tiempo medido, pero si por alguna razón lo quieren antes eso puede pegar en los planes del equipo, solo eso, en general es un buen equipo todo se ayudan y es un buen ambiente laboral.

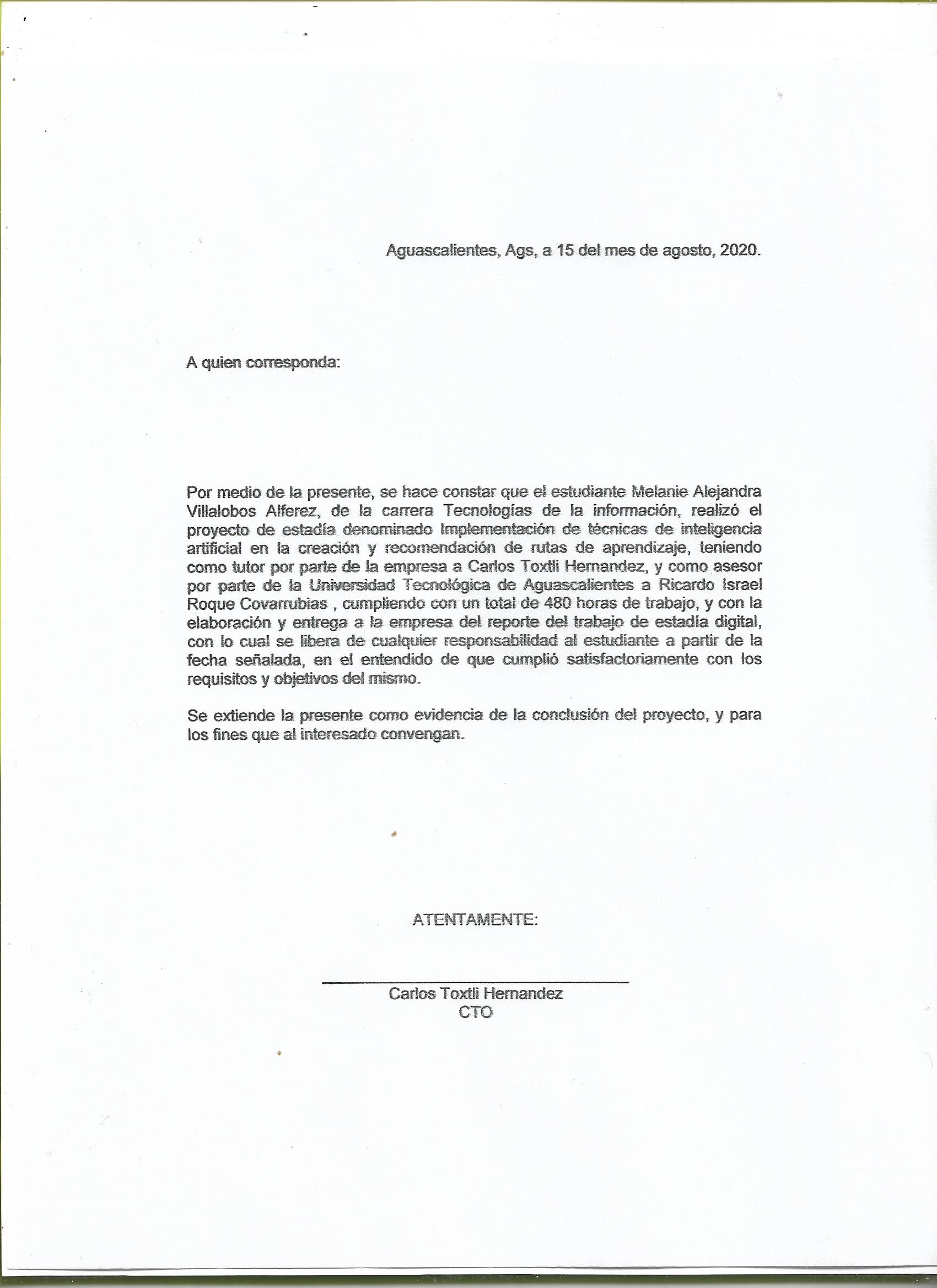
## Anexos.

**Cronograma de actividades.**

****

**Carta de Liberación en hoja membretada**

(En caso de que la empresa no cuente con hoja membretada, la carta deberá contar con sello de la empresa).



**Formatos varios**

## Referencias Bibliográficas

Juan Ignacio Bagnato. (November 14, 2017). Aprende Machine Learning. 25/07/2020, de Proudly powered by WordPress Sitio web: <https://www.aprendemachinelearning.com/que-es-machine-learning/>

G&L Group S.A. Monteagudo. (2020). Diferencia entre Machine Learning y estadísticas. 10/07/2020, de Distrito Tecnológico Sitio web: <https://www.gylgroup.com/novedades/diferencia-machine-learning-estadisticas>

Mikel Niño e. (10 de noviembre de 2016). Conceptos básicos de “Machine Learning” (Aprendizaje Automático): algoritmos vs modelos. 28/07/2020, de InShare Sitio web: <http://www.mikelnino.com/2016/11/conceptos-basicos-machine-learning-aprendizaje-automatico-algoritmos-modelos.html>

DEAL. (diciembre 20, 2007). Laboratorios de Innovación. 29/07/2020, de DEAL Sitio web: <https://www.deal.com.br/tecnologia-y-innovacion/laboratorios-de-innovacion-iot-ai-machine-learning/?lang=es>