Tabla de Contenido

[Introducción 4](#_Toc2966858)

[Descripción del problema 4](#_Toc2966859)

[Alcance 4](#_Toc2966860)

[Justificación 4](#_Toc2966861)

[Estado del arte 4](#_Toc2966862)

[Aplicaciones usando Machine Learning usados en eSports 4](#_Toc2966863)

[Objetivos 4](#_Toc2966864)

[Objetivo General 4](#_Toc2966865)

[Objetivos Específicos 4](#_Toc2966866)

[Metodología 4](#_Toc2966867)

[Marco teórico 5](#_Toc2966868)

[Machine Learning 5](#_Toc2966869)

[Conceptos generales sobre MOBA 5](#_Toc2966870)

[Partidas individuales de jugadores profesionales en Dota 2 5](#_Toc2966871)

[Partidas equipos profesionales en Dota 2 5](#_Toc2966872)

[Algoritmos de predicción usados en e-Sports 5](#_Toc2966873)

[Recolección y Análisis del conjunto de datos 5](#_Toc2966874)

[Recolección de datos usando Rest API 5](#_Toc2966875)

[Análisis de datos 5](#_Toc2966876)

[Identificación de variables 5](#_Toc2966877)

[Aplicación de algoritmo, experimentos y análisis de partidas individuales 5](#_Toc2966878)

[Aplicación de redes neuronales 5](#_Toc2966879)

[Experimentos y detección de variables dependientes 5](#_Toc2966880)

[Análisis de resultados y hallazgos 5](#_Toc2966881)

[Aplicación de algoritmo, experimentos y análisis de equipos profesionales 6](#_Toc2966882)

[Aplicación Ranfom Forest 6](#_Toc2966883)

[Experimentos y detección de variables dependientes 6](#_Toc2966884)

[Experimentos Usando variable dependiente aplicando resultados de partidas individuales 6](#_Toc2966885)

[Análisis de resultados y hallazgos 6](#_Toc2966886)

[Pruebas 6](#_Toc2966887)

[Trabajo Futuro 6](#_Toc2966888)

[Conclusiones 6](#_Toc2966889)

[Referencia o Bibliografía 6](#_Toc2966890)

[Apéndice 6](#_Toc2966891)

# Introducción

## Descripción del problema

## Alcance

Los entregables de esta investigación son se describen a continuacion:

### API

Se entregará un API o Application Programming Intergace, que servirá para hacer uso de los objetivos de este proyecto de investigación, el cual cumplirá con un conjunto de características descritas en la siguiente sección.

### Características principales

Característica #1: Consiste en que el sistema se alimentará de los equipos profesionales en el contexto del juego Dota 2. Su salida devolverá que equipo será el ganador del torneo a predecir.

Característica #2: Se ingresará el ID de 10 jugadores, al menos uno profesional. El sistema retornará los jugadores profesionales que ganen la partida.

Característica #3: Se ingresará 2 equipos, el sistema retornará cual equipo gana o pierde la partida.

### Documentación API

El API se entregará con una documentación fuera del código, que describirá cada una de las funciones relevantes que se llevaron a cabo para esta investigación. Listando los diferentes scripts y algoritmos para lograr la predicción de las diferentes características y así mismo su debido uso.

También, se entregará un archivo README, explicando como se debe ejecutar el uso de este API, con sus debidos requerimientos de hardware y requisitos de software.

### Registro de Software ante entidades competentes.

(no se si esto en realidad sea necesario o requerido)

El API o software para este proyecto se entregará registrado ante el organismo de control de Colombia, que en este caso es La Oficina de Registro de la Dirección Nacional de Derecho de Autor.

Aquí seria bueno hablar sobre los entregables, el software o casos de usos que serán el resultado.

Posibles casos de Uso:

Entrada: equipos participantes en un torneo -> salida el ganador. (ppal)

Entrada: jugador -> salida gana o pierde

Entrada: Equipo participante -> salida: gana o pierde

## Justificación

### Descripcion del problema

El estudio en el campo de los deportes electronicos o e-sports de aquí en adelante cobra cada dia mas relevancia dado que cada vez son mas los millones de personas que alrededor del mundo usan este tipo de deportes. En este campo no existen estudios solidos que nos permitan propones modelos estrategicos, debido a que existen cientos de juegos alrededor del mundo, cada vez mas complejos debido a lo sofisticados que estos pueden llegar a ser.

Es el caso de Dota 2, un juego electronico que a su vez se ha convertido en un deporte electronico que ha venido evolucionando desde su año de aparicion 2013. Dota 2, es un juego de tipo Multiplayer Online Mobile Arena o MOBA de aquí en adelante.

En cuanto a Dota 2 como un deporte electronico, desde su inicio contó con una secuencia de ligas que reune los jugadores mas hábiles para llevarse consigo un premio millonario. En la actualidad, se desarrollan varias ligas y torneos internacionales cada vez con mas interés de diferentes empresas que a su vez invierten en el campo de los e-sports.

En los ultimos 7 años se vienen desarrollando The Intenational, que es evento mas importante internacional de este deporte. Es un torneo organizado por Valve, que a través de los años ha venido surgiendo e incrementado los precios para el ganador. Para el torneo del The International 2018, el total premio para los equipos ganadores seria de un total de 25 millones de dolares. [liquipedia]

Sin embargo, estos torneos están sufriendo grandes cambios en rotacion de jugadores, equipos mas competentes debido a las tecnologias usadas para entrenar y cambio constante de estrategias para asumir estos campeonatos.

En la actualidad, ninguno de los proyectos investigativos se ha formulado alguna metodologia o API, ayudar a los equipos que compiten en este tipo de certamen y que estos equipos se puedan apoyar las de estrategias teniendo en cuenta factores que ocurren en los campeonatos como lo es la rotacion de jugadores en cada equipo. Asi como en un equipo de futbol, los equipos en Dota 2 rotan cada cierto tiempo de jugadores para sorprender a sus equipos rivales con nuevas estrategias.

Tomando como base la investigacion realizada por [Result Prediction by Mining Replays in Dota 2], se pretende realizar un modelo predictivo para equipos profesionales en Dota 2 teniendo en cuenta el rendimiento que tiene cada jugador profesional y este a su vez sirva como entrada para determinar si este equipo profesional ganará o no un torneo.

### Justificación

Estudios relacionados en e-sports, sugieren por lo general modelos predictivos en partidas contra Bots [paper adaboost], modelos predictivos usando arboles de decisión implementando Random Forest [RF data mining], algunos otros investigan jugadores generalmente de cualquier tipo, profesional o amateur, hasta ahora ninguno hace una investigacion que tenga en cuenta los equipos profesionales y su participacion en los torneos. Destacando esta falta de investigacion y con animo de proponer una manera de estrategia para algun equipo profesional, se prentende proponer un nuevo modelo predictivo que se basa en la eficiencia de predecir el ganador de un jugador profesional de un partida tipo rank.

Una de las carescteristicas por las que cualquier jugador profesional o amateur atraviesa en Dota 2, es que todos compiten por alcanzar el primer lugar de su region, la forma con que Valve a generado esta tabla es a través de su Matchmaking Rating o MMR de aquí en adelante.

Antes de entrar al detalle de como eventualmente un jugador gana o pierde puntos, vale aclarar que el sistema de Matchmaking exige que un jugador no puede ganar del 100% de partidas el total de esas partidas. El sistema en la asignacion de equipos al momento de busqueda se asegurará de que el jugador tenga 50% de probabilidades de ganar una partida. Por lo tanto, el sistema asigna a susc compañeros de equipo para que ganen y pierdan 50/50 [https://dota2.gamepedia.com/Matchmaking\_Rating].

El MMR se gana a través de juegos Ranked, donde cualquier tipo de jugador gana o pierde puntos en. cada partida, en la actualidad no existe ninguna tabla que exactamente muestra las reglas sobre bajo que condiciones cuantos puntos se puede ganar puntos o pueda perder puntos, pero existen ciertas reglas que hacen que los jugadores entiendan cuanto podria sumar o restar cada partida. Veamos algunas:

El MMR se puede calcular jugando Solor en modo ranked, o se puede calcular jugando en Equipo (party) o con amigos en modo Ranked. Los modos Ranked pueden ser Modo Capitan, Modo Normal y modo aleatorio.

Bajo la condicion solo en modo ranked, si el jugador compite en Modo Normal y gana, podria obtener 30 puntos, sin embargo esto dependerá de que porcentaje de ganancia le asigna al jugador el sistema teniendo en cuenta su actual tasa de gana/pierda. Si tiene menos porcentaje de ganar, y gana la partida bajo esa condicion ganará mas puntos que si Gana una partida con un alto porcentaje de Ganar.

Este mismo jugador con el mismo porcentaje de ganar, pero jugando con amigos (party) en modo ranked, el sistema calculará el porcentaje de cada uno de los miembros del equipo y hará un calculo gana/pierde para determinar un equipo oponente, además teniendo en cuento el MMR del equipo en total, para encontrar un equipo opuesto. Bajo esta condicion si gana este jugador la partida, no ganaria los mismos puntos MMR como en el caso anterior.

En conclusion, es complejo tener una compilacion de todos los puntos o reglas ya que no existen de manera publica en algun recurso publico de Valve o paginas de Dota 2, pero nos da a entender que el sistema de MMR afecta a cualquier tipo de jugador en Dota 2 y que no es facil llegar a estar en la cima de la tabla de posicionamiento de cada region. Es natural que un jugador que participe en una liga profesional se encuentre entre los primeros 100 jugadores de la region en la que esté asignado según geograficamente viva. Por ultimo, cada parche de juego o actualizacion podria cambiar estas reglas las cuales son dadas a conocer en cada nota de la actualizacion, lo cual podria hacer que para un jugador sea mas complejo subir puntos.

### Estado del arte

Aquí quiero escribir sobre la motivación, en la actualidad me encuentro hablando con un equipo de Sudamérica que esta constantemente participando en algunos torneos del continente y ocasionalmente contra otros equipos de otros continentes. Al momento no hay nada pactado pero si sale me gustaría usarlos como justificación y motivación aparte que quiero hacerlo para apuestas $.

Vale la pena hablar aquí de la línea del tiempo ? Crecimiento o ganancias obtenidas en el mundo de los eSports? En ese orden no de si la descripción del problema venga después de esta Justificacion. Y si es así, vale la pena tener entonces un apartado solamente de esto.

## Estado del arte

Aquí quiero hablar de otras tesis y artículos que me pueden ayudar en esta investigacion

### Aplicaciones usando Machine Learning usados en eSports

Aquí quiero hablar de lo que se ha hecho y algoritmos que se han usado par apredecir el ganador de una partida. Prácticamente quiero hablar de las otras tesis que ya han usado ANN para predecir y RF , de hecho es por eso que voy a usar ANN y RF, quiero usarlas ya que otros ya han hecho algo al respecto en el estado del arte.

# Objetivos

Los objetivos son la pieza clave, en esta sección presento el objetivo general y los objetivos específicos.

## Objetivo General

Diseñar un modelo predictivo usando máquinas de aprendizaje supervisado que prediga el equipo ganador en un torneo o liga de Dota 2 dado los datos parciales recopilados de ligas anteriores y partidas comunes de cada jugador profesional que participe en estas ligas

## Objetivos Específicos

* Determinar cuál es la relación entre las partidas que juega un jugador profesional fuera de la liga y el rendimiento de ese jugador cuando participa en una liga profesional.
* Proponer un modelo predictivo de que un equipo pueda o no ganar una partida en una liga teniendo en cuenta el historial de ligas.
* Elaborar un modelo en el cual se analicen los resultados singulares de los jugadores profesionales en sus partidas fuera de campeonato.

# Metodología

Es necesario dedicar una sección a metodología , si implícitamente se está siguiendo la metodología a través del documento.

# Marco teórico

## Machine Learning

Machine learning se puede definir como los métodos computacionales que usan experiencias conocidas para mejorar el rendimiento o hacer predicciones precisas [Foundations of Machine Learning by M. Mohri, A. Rostamizadeh, A. Talwalkar]. Donde la experiencia se expresa como la información pasada que está disponible para aquel sujeto que aprende, en este caso la maquina. Estos datos que son etiquetados a través de las interacciones con la vida real se encuentran reunidos en forma de colección de datos para su posterior procesamiento.

Machine learning consiste en diseñar algoritmos de predicción eficientes y precisos basados en un modelo matemático conocidos como datos de entrenamiento, en orden de hacer predicciones o decidir sin ser explícitamente programados para realizar una tarea [Bishop, C. M. (2006), Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2].

## Artificial Neural Network - ANN

Una red neuronal artificial es un modelo computacional inspirado en la red neuronal cerebral. Esta compuesta por un conjunto de neuronas las cuales son las unidades de procesamiento, que se encuentras interconectadas con otra red de neuronas. La unión entre esta rede neuronas tiene un peso y este peso determinan el impacto de una neurona a otra [1S.V. Kozyrev,” Classification by ensembles of neural networks”, Steklov Mathematical Institute, February 21, 2012.].

## Random Forest

Vale la pena hacer una introducción en machine learning, tal vez algunos algoritmos que hacen o cuales es por ejemplo la motivación de ANN y de RF.

## Conceptos generales sobre Dota 2

## Dota 2

Dota 2 es un Multiplayer Online battle arena o MOBA de aquí en adelante, en donde dos equipos de 5 jugadores contra 5 jugadores compiten en un mapa para destruir el edificio principal llamado Ancient (por eso el nombre Defense Of The Ancients) del equipo contrario. El equipo que primero logre destruir tal edificio gana la partida. El juego tiene disposicion un total de 115 personajes o Heroes como son llamados en el juego, cada heroe tiene diferentes caracteristicas [http://www.dota2.com/heroes].

Uno de los componentes claves que se encuentra en Dota 2, es la estructura organizacional que se lleva a cabo en los torneos o ligas que dan origen en el año 2011, donde docenas de equipos compiten entre si para llevarse actualmente alrededor de 11 millones de dolares para el equipo ganador [https://liquipedia.net/dota2/The\_International/2018].

Para que un equipo pueda llegar a uno de estos certamen internacional existen dos manera que se puede lograr, uno es a través de una invitación que Valve expide cuando un equipo logra ciertos puntos minimos, una vez logrado clasifican o por otro lado son invitados para participar en una liga profesional.

La otra manera consiste en un campeonato que se organiza por lo general de manera virtual, sin embargo miles de equipos por region compiten en modo brackets, por lo general por region, digamos Europa compite por un cupo, sur america tiene un cupo, asia medio tendria otro cupo y asi, el que gane este campeonato se gana el cupo para el campeonato internacional.

De esta manera se logra la participacion, para el caso del The International 2019 de reunir 18 equipos compuestos por 12 equipos invitados a través de puntos por liga y 6 equipos por region calificados por medio de brackets por campeonato virtual [https://liquipedia.net/dota2/The\_International/2019].

Quiero hablar sobre la importancia de MOBA y tal vez de los eSports , no se si deba mencionar en el titulo en vez de MOBA – eSports.

### Partidas individuales de jugadores profesionales en Dota 2

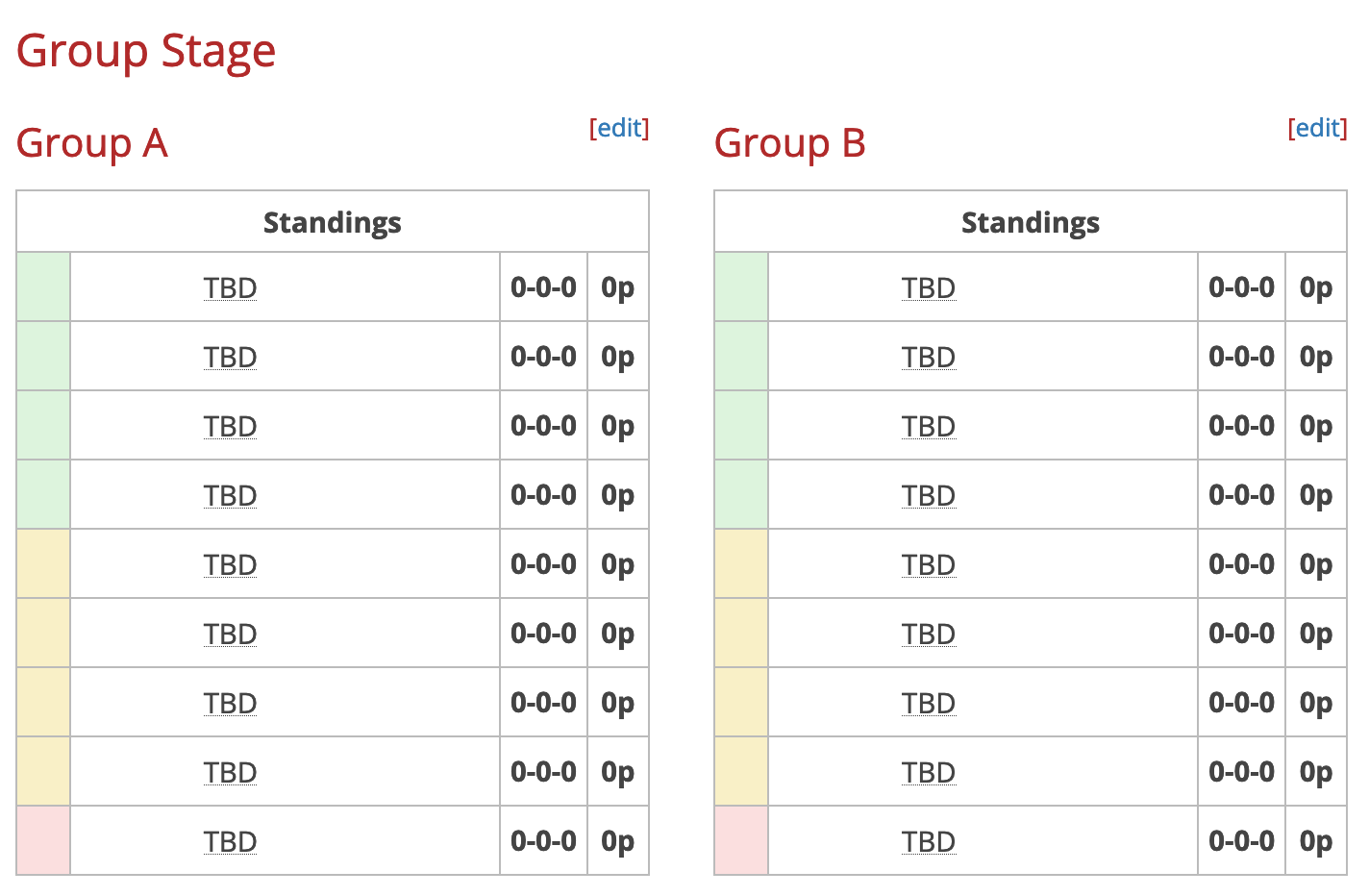
La partidas individuales que se explicarán a continuación, serán aquellas que suben el MMR del jugador y serán esta las partidas collecionadas en la base de datos para un pre procesamiento para un posterior diseño del modelo predictivo.

* *Juego Normal*: Se compone de 5 jugadores contra 5 jugadores, en al cual cada uno de los jugadores por turnos, eligirá el heroe que desea jugar durante el juego. En el inicio de esta selección de heroes, cada jugador podrá prohibir que algunos heroes puedan ser elegidos y para esto tienen 30 segundos. Los heroes prohibidos aplican para los 2 equipos participantes.
* *Modo Capitan*: El modo capitan se basa en que 1 solo jugador elegido por el sistema con el MMR mas alto de cada equipo. Este jugador o capitan tendrá la responsaiblidad de elegir los heroes de cada uno de sus integrantes por turnos. En cada turno podrá prohibir 2 heroes al equipo contrario bajo un orden de tiempo total. El tiempo que tiene cada capitan en total son 130 segundos para Elegir, y puede usar ese tiempo como desee. Por ultimo, el capitan es elegido por el sistema como aquel que tiene el MMR mas alto. [https://dota2.gamepedia.com/Game\_modes]
* *Modo Aleatorio*: El sistema aleatoriamente genera 50 heroes unicamente y todos los jugadores tienen hasta 20 segundos para elegir un heroe por medio de turnos.

Explicar aquí las partidas individuales, como se componen, datos generales de una partida, no existe diferencia entre una partida normal a una partida de un jugador profesional.

### Partidas equipos profesionales en Dota 2

En los torneos o ligas profesionales, los tipos de partida son los de tipo Modo Capitan. Por lo general en estas ligas el formato de competencia se basa en Round Robin, los equipos son divididos en 2 grupos de 9 equipos, donde cada uno de los equipos compite contra los 8 equipos del grupo. Cada equipo jugará 2 veces contra cada equipo en el grupo asignado. Cada victoria sera 1 punto. Si un equipo pierde 2 partidas contra un equipo no tendrá en esta etapa ningun valor. Solo serán clasificados a través de los puntos conseguidos.



[https://liquipedia.net/dota2/The\_International/2019]

Al final de esta etapa habrán 4 equipos en cada grupo que serán los equipos que iniciarán en la parte arriba de los brackets, los 4 de abajo serán parte del bracket de abajo, y los 2 ultimos equipos de cada grupo serán inmediatamente eliminados.



Los equipos localizados en upper bracket podran jugar 3 partidas contra otro equipo de los upper bracket, si uno de estos equipos pierde 2 juegos, pasa inmediatamente a Lower bracket. El ganador pasará al siguiente round.

Los equipos localizados en lower bracket, solo podrán jugar 1 partida contra un equipo, si pierden esta partida quedarán inmediatamente eliminados en la primera etapa. El ganador aun tendrá oportunidad y pasara a la siguiente etapa en lower bracket.

En el round dos, los equipos de lower bracket, podran volver a jugar 3 partidas al igual que los de upper bracket. Cualquier equipo que pierda en lower bracket será eliminado.

Al final de esta serie de rounds, quedarán 2 equipos que competirán por el primer lugar de la liga, y solo en este caso podrán jugar 5 partidas para llevarse el premio. Ganará aquel equipo que logre 3 victorias en este ultimo round.

## Algoritmos de predicción usados en e-Sports

Support Vector Machine y Artificial Neural Network: [Predicting outcomes of professional DotA 2 matches] Experimenta con dos modelos predictivos teniendo cuenta las variables ‘gold per minute’, ‘experience per minute’, ‘kills per minute’, ‘lane efficency’ y ‘solo competitive rank’. En estos experimentos los autores implementan Support Vector Machine con Kernel Lineal y Redes Neuronales comparando el set de entrenamiento de los dos algoritmos, concluyen finalmente que SVM tiene la mayor exactitud.

Naïve Bayes con Ada Boost: [Outcome Prediction of DOTA2 Based on Naive Bayes Classifier] Los autores entrenado un data set que se encontraba disponible en UCI Machine Learning Repository, demostrando que este algoritmo sirve como herramienta para predecir que un equipo puede o no ganar una partida.

Dado el estado de arte que algoritmos de predicción se han usado para eSports o mas bien específicamente para Dota 2

# Recolección y Análisis del conjunto de datos

## Recolección de datos usando Rest API

Aquí quiero hablar sobre Open Dota y la necesidad de construir un software que haga el llamado masivo de datos. Pregunta: debería hablar de la tecnología usada, la necesidad de tener una base de datos y los llamados masivos en diferentes tiempos para lograr obtener los datos deseados ?

## Análisis de datos esto en en un anexo en una tabla y descripción de los datos en un anexo

Simplemente hablar del JSON que se obtiene, relevancia. ¿Qué otra cosa?

**Open dota** es un repositorio privado de una empresa que llamam GOSU, proveen una gran variedad en su API, donde podremos encontrar partidas de jugadores profesionales, resultados de ligas profesionales, lista de jugadores profesionales entre otros.

Esta investigacion se centrará principalmente en dos llamados. Partidas de jugadores profesionales y Resultados generales de las ligas profesionales.

Partidas de jugadores profesionales /matches/{match\_id}

Para lograr este set, fue necesario hacer un primer compilado de los equipos y asi mismo de los jugadores que componen cada equipo, para lograr este masivo conjunto de datos fue importante tener en una collecion los Ids de partidas jugados por cada uno de los jugadores.

Finalmente despues de tener el compilado de los Ids de partidas se logro iterar sobre el API de partidas individuales para descargar todo el ser y almacenarlo en una base de datos. De aquí nace una gran problemática que ese el espacio en disco y procesamiento que esto exige, estamos hablando de una tabla que puede llegar a medir 15 GB. Lograr procesas este gran dataset exige diseñar, construir e implementar una solucion que pueda procesar este volumen en maquinas locales.

Aquí las variables dependientes e independiente

Resultados de ligas profesionales. /teams/{team\_id}/matches

Aquí las variables dependientes e independiente

### Identificación de variables

Aquí me gustaría hablar sobre las variables que considero que serian un buen inicio, no se si valga la pena o profundizar en la identificación de variables hasta que haga los experimentos en la siguiente sección.

# Aplicación de algoritmo, experimentos y análisis de partidas individuales

## Aplicación de redes neuronales

Aplicación me refiero a aplicar el algoritmo de ANN en partidas individuales, debo de ser tan específicos en los títulos ¿? Lo mismo para equipos profesionales.

## Experimentos y detección de variables dependientes

## Análisis de resultados y hallazgos

# Aplicación de algoritmo, experimentos y análisis de equipos profesionales

## Aplicación Ranfom Forest

## Experimentos y detección de variables dependientes

### Experimentos Usando variable dependiente aplicando resultados de partidas individuales

## Análisis de resultados y hallazgos

# Software

Vale la pena hablar del software que se va a entregar? En termino de uso y registro?

# Trabajo Futuro

# Conclusiones

# Referencia o Bibliografía

# Apéndice