Entregable 1

# Título del tema de tesis

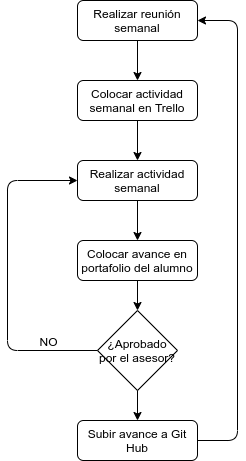
Creación de robots conversacionales inteligentes basados en Machine Learning.

# Asesor

El asesor para este tema de tesis es el profesor Cesar Beltrán Castañon.

La metodología de trabajo será la siguiente: Se tendrán reuniones virtuales, presenciales cuando sea posible, en las cuales se discutirá sobre los temas avanzados por el alumno. En dichas reuniones, el alumno presentará al profesor su avance semanal y también se discutirá sobre los próximos avances a realizar. Se dispondrá de un repositorio en GitHub en donde el alumno subirá sus actividades realizadas, código de la implementación, así como también los avances revisados, aprobados y corregidos. Se piensa usar Trello para mantener un control de las actividades. Cabe resaltar que el repositorio de GitHub será usado en conjunto con la portafolio del alumno será creado por los profesores del curso, como se menciona en el documento de plan de tesis.

En la siguiente imagen se presenta un flujograma de la metodología de trabajo.



El horario de reuniones acordado con el asesor serán los martes a las 9 de la mañana. La reunión durará al menos 1 hora.

# Área

Este proyecto pertenece al área de Ciencia de la Computación y a la rama de Procesamiento de lenguaje Natural.

# Problemática

Las personas de distintas edades y de distinto contexto no necesariamente poseen el la misma capacidad para poder entender un texto, ya sea hablado o escrito. Como primer ejemplo, se tiene el caso de que un alumno de primaria no necesariamente tiene la capacidad de entender los textos de un alumno de secundaria. Como segundo ejemplo se puede tomar el caso de un médico y un paciente que necesitan saber sobre un mismo tema. Es muy probable que el médico entienda términos más complejos que el paciente. Entonces, con estos ejemplos, se descubre que existe la necesidad de que los robots conversacionales sean entrenados con textos de distinta complejidad para que puedan dar una respuesta adecuada, con respecto a la complejidad de la respuesta, a los usuarios. La pregunta que se busca resolver en la problemática es: ¿Como identificar el nivel de complejidad de un texto referente a un tópico?

# Resultados esperados

Al finalizar este proyecto de tesis, se espera disponer con una libería en Python que permita al usuario etiquetar textos en español según su complejidad. Inicialmente se tiene pensado que la complejidad obtenida sea un valor numérico. Por ejemplo, nivel 1, nivel 2, nivel 3, nivel 4 y nivel 5. Esto permitirá al usuario poder etiquetar textos para que los pueda usar en el entrenamiento de robots conversacionales.

# Métodos y procedimientos

Para el desarrollo de esta tesis se usarán las siguientes herramientas:

* GIT: Esta herramienta de control de versiones será usara para versionar el código del modelo.
* Visual Studio Code: Este editor de texto será usado para escribir el código del modelo.
* Google Colab: Esta herramienta online será usada para crear documentos de “Jupyter Notebook” para probar el código, mostrar gráficos y resultados.
* Jupyter Lab: Herramienta que será usada para crear documentos de “Jupyter Notebook” en donde probar el código, mostrar gráficos y resultados. Será usado cuando las limitaciones de Google Colab no sean suficientes.
* Python: Python: Lenguaje de programación multiparadigma en el que será desarrollada el modelo.
* FreeLing: Es una Suite de herramientas de analisis de lenguaje de código abierto que puede ser utilizada para complementar el desarrollo del modelo.