



UNIVERSIDAD DEL BÍO-BÍO
FACULTAD CIENCIAS EMPRESARIALES

SOLICITUD INSCRIPCIÓN ACTIVIDAD DE TITULACIÓN

1. Identificación Alumno (S).

Nombre	: Francisco Fernando Muñoz Coletti
Dirección	: O'Higgins 1500 Quillón
Teléfono	: +56 9 4290 8358
E-Mail	: franmucol16@gmail.com
Carrera	: Ingeniería Civil Informática
Depto.	: Sistemas de Información

2. Evaluadores

2.1 Profesor guía

Nombre : Pedro Gerónimo Campos Soto	Firma:
-------------------------------------	--------

2.2 Profesor co-guía

Nombre : Nathalie Risso Sepúlveda	Firma:
-----------------------------------	--------

3. Proyecto

3.1 Título Que Identificará La Actividad De Titulación.

Análisis de comentarios usuarios de la red social Twitter en temáticas de contingencia nacional en territorios específicos de Chile.

3.2 Objetivos Generales Y Específicos De La Actividad De Titulación.

3.2.1 Objetivos General

Categorizar opiniones de usuarios de Twitter en territorios específicos de Chile sobre temas de interés nacional mediante Análisis de Sentimientos.

3.2.1 Objetivos Específicos

- 1 Revisar literatura relacionada a análisis sentimental de textos en redes sociales, uso de la API de Twitter y geolocalización en formato GeoJSON.
- 2 Identificar y filtrar usuarios y sus respectivos Tweets pertenecientes a territorios específicos dentro de Chile mediante las funcionalidades de herramientas como Tweepy y GeoJSON
- 3 Seleccionar y entrenar un modelo de análisis de sentimientos para aplicar a temas de interés nacional sobre Tweets obtenidos.
- 4 Implementar una interfaz gráfica sencilla para visualizar los resultados obtenidos.

3.3 Justificación Del Proyecto Propuesto.

Las redes sociales en la actualidad poseen un muy alto flujo de información, sea formal o informal. Alrededor del 60.1% de la población mundial actualmente utiliza internet (4.7 billones de personas). De estos usuarios, el 92.12% utiliza plataformas de redes sociales activamente. (Kemp, 2021). Si se observa exclusivamente en Chile, un 83.5% de su población (16 Millones) son usuarios activos en redes sociales (Alvino, 2021).

El uso de dichas redes va desde el ocio, pasando por columnas de opinión hasta incluso denuncias ciudadanas sobre problemáticas que se viven día a día en nuestro país. Estas plataformas se han convertido actualmente en la voz del pueblo chileno y su uso se intensificó

aún más post 18 de Octubre de 2019 (en comparación a meses anteriores) (Ferreira, 2020). Sin embargo esta información queda registrada, mas no organizada para el posterior estudio de la misma por las dificultades que se presentan a la hora de clasificarla: Cuentas fraudulentas, información falsa, el alto costo que significaría tener un equipo encargado de recolectar y analizar los sentimientos expresados en dicha información de forma análoga, entre otras.

Es por todo lo anteriormente mencionado que en este proyecto de tesis se propone: Desarrollar un sistema automatizado de análisis de sentimientos destinado a clasificar las opiniones de usuarios de la red social Twitter con respecto a temas de interés nacional y social acorde a territorios específicos del país. Adicionalmente esto permitiría facilitar y reducir costos a empresas dedicadas a la investigación de opinión (por ejemplo: CADEM, IPSOS, CEP, Etc.)

3.4 Actividades a realizar

- 1 Revisar literatura sobre análisis de textos en redes sociales, manejo de la API de Twitter mediante Tweepy, obtención de lugares geográficos mediante formato GeoJSON, Opinion Mining, Python 3 y Deep Learning.
- 2 Recopilar datos requeridos desde API de Twitter (Streaming y Rest) y almacenarlos en una base de datos local.
- 3 Definir criterios reglas y palabras para filtrar a los potenciales usuarios fraudulentos, cuentas de entidades y/o Tweets no útiles para este estudio.
- 4 Desarrollar módulo para identificar el territorio al cual pertenecen los Usuarios emisores de los Tweets.
- 5 Seleccionar y entrenar un modelo de Análisis de Sentimientos.
- 6 Realizar experimentos con la finalidad de evaluar el modelo ya entrenado y almacenar dichos resultados para su posterior representación.
- 7 Graficar los resultados obtenidos posterior al procesamiento realizado.
- 8 Desarrollo de Informe y Presentación para la defensa de la Tesis.

3.5 Plan De Trabajo A Desarrollar

Actividad	Rango de fechas
Revisar literatura sobre análisis de textos en redes sociales, manejo de la API de Twitter mediante Tweepy, obtención de lugares geográficos mediante formato GeoJSON, Opinion Mining, Python 3 y Deep Learning.	03/09/2021 - 20/10/2021
Recopilar datos requeridos desde API de Twitter (Streaming y Rest) y almacenarlos en una base de datos local.	15/09/2021 - 15/10/2021
Definir criterios, reglas y palabras para filtrar a los potenciales usuarios fraudulentos, cuentas de entidades y/o Tweets no útiles para este estudio.	16/10/2021 - 26/10/2021
Desarrollar módulo para identificar el territorio al cual pertenecen los Usuarios emisores de los Tweets.	27/10/2021 - 11/11/2021
Seleccionar y entrenar un modelo de Análisis de Sentimientos.	12/11/2021 – 16/01/2022 (Aprox 9 semanas, dejando enero y febrero como vacaciones)
Realizar experimentos con la finalidad de evaluar el modelo ya entrenado y almacenar dichos resultados para su posterior representación.	15/03/2022 - 15/04/2022
Graficar los resultados obtenidos posterior al procesamiento realizado.	16/04/2022 - 16/05/2022
Desarrollo de Informe y Presentación para la defensa de la Tesis.	17/05/2022 - 17/06/2022

3.6 Trabajos Similares Realizados Previamente.

Ferreira Sandoval, M. Análisis de cambios en comentarios de usuarios de la red social Twitter a partir de la crisis social en Chile. Septiembre, 2020. En dicho trabajo se realizó un estudio de actividad de la red social Twitter en Chile y como cambió esta durante 3 meses antes y 3 meses después del 18 de octubre de 2019.

A diferencia del trabajo antes mencionado, este proyecto de tesis tiene como objetivo:
1.- Utilizar herramientas de geolocalización para obtener datos de los usuarios pertenecientes a

territorios específicos del territorio nacional.

2.- Obtener opiniones respecto a dos temáticas de importancia nacional como lo son la salud y las pensiones. Definir un listado de palabras para entrenar el modelo de análisis de sentimientos para luego realizar la fase de experimentación y finalmente representar gráficamente los resultados.

3.7 Descripción De Los Aspectos Fundamentales De La Metodología A Utilizar (Software)

Este proyecto genera la necesidad de tener una alta cantidad de filtros, reglas y pruebas para que el estudio sea lo más preciso posible. Es por esto que se utilizará la metodología espiral para el desarrollo del Software, donde se intenta reducir al mínimo los fallos, realizar constantemente una planificación y gestión de estos para definir cuales presentan más riesgos y buscar estrategias para lidiar con ellos. Al final de cada iteración del proceso espiral se generará un bloque entregable el cual será evaluado en conjunto al profesor guía. Una vez el software esté en las etapas finales, se realizará una última evaluación antes de dejar el software operativo y listo para pasar a la fase de mantención.

4. Bibliografía

4.1 Referencia bibliográfica


- 1 Ferreira Sandoval, M. (Septiembre, 2020). Análisis de cambios en comentarios de usuarios de la red social Twitter a partir de la crisis social en Chile.
- 2 Vinodhini, G., & Chandrasekaran, R.M. (2012). Sentiment Analysis and Opinion Mining: A Survey(6ªed., Vol.2).
- 3 Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S., Schaub, T., & Wilde, E. (2016). *The GeoJSON Format*.
- 4 Chollet, F. (2017). *Deep Learning with Python*. Manning Publications.

4.2 Linkografía

- 1 Brownlee, J. (2019, 16 septiembre). *A Gentle Introduction to Transfer Learning for Deep Learning*. Machine Learning Mastery. Recuperado el 12 de septiembre de 2021, de <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- 2 Education, I. C. (2021, 12 agosto). *Machine Learning*. IBM. Recuperado el 12 de septiembre de 2021, de <https://www.ibm.com/cloud/learn/machine-learning>

- 3 Twitter, Inc. (s. f.). *Getting Started with Our Platform*. Docs | Twitter Developer Platform. Recuperado 8 de septiembre de 2021, de <https://developer.twitter.com/en/docs/getting-started>
- 4 Kemp, S. (2021, 21 abril). *60 percent of the world's population is now online*. We Are Social. <https://wearesocial.com/blog/2021/04/60-percent-of-the-worlds-population-is-now-online>
- 5 Alvino, C. (2021, 3 junio). *Estadísticas de la situación digital de Chile en el 2020–2021*. Branch. <https://branch.com.co/marketing-digital/estadisticas-de-la-situacion-digital-de-chile-en-el-2020-2021/>

LA PRESENTE SOLICITUD DE INSCRIPCIÓN DE LA ACTIVIDAD DE TITULACIÓN
SIGNIFICA UN COMPROMISO DE
CUMPLIR LO ESTIPULADO EN ELLA.



FIRMA ALUMNO

FECHA PRESENTACIÓN SOLICITUD: 6 de Octubre de 2021

RESOLUCIÓN DIRECTOR DE ESCUELA/ JEFE DE CARRERA

APROBADO

☐

REPROBADO

☐

OBSERVACIONES:
.....
.....
.....

Linkografía

FIRMA DIRECTOR DE ESCUELA

FECHA RESOLUCIÓN: