

UNIVERSIDAD PERUANA UNIÓN
FACULTAD DE INGENIERIA Y ARQUITECTURA
Escuela Profesional de Ingeniería de Sistemas



**Análisis de sentimientos en Twitter con stacking ensemble
para el índice de aprobación política de la presidencia del
estado peruano**

Tesis para obtener el Título Profesional de Ingeniero de Sistemas

Autor:

Bach. Deivis Ronald García Cercado
Bach. Juan Manuel Villarroel López
Bach. William Salvadory Barzola Torres

Asesor:

Dr. Juan Jesús Soria Quijaite

Lima, 06 de noviembre de 2024

DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo Dr. Soria Quijaite Juan Jesús, docente de la Facultad de Ingeniería y Arquitectura, de la Escuela Profesional de Ingeniería de Sistemas, de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: “**Análisis de sentimientos en Twitter con stacking ensemble para el índice de aprobación política de la presidencia del estado peruano**” de los autores García Cercado, Devis Ronald, Villarroel López, Juan Manuel y Barzola Torres, William Salvador tiene un índice de similitud de 8% verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima a los 8 días del mes de enero del año 2025.



Dr. Soria Quijaite Juan Jesús

ACTA DE SUSTENTACIÓN DE TESIS



ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Ñaña, Villa Unión, a 06 día(s) del mes de noviembre del año 2024 siendo las 09:00 horas, se reunieron los miembros del jurado en la Universidad Peruana Unión Campus Lima, bajo la dirección del (de la) presidente(a):

Mg. Geraldine Verónica Alvizuri Ureña el (la) secretario(a): Mg. Lenin Henry Centurión Julca
 y los demás miembros: Mg. Nemias Saboya Ríos
Mg. Ferdinan Edgardo Pineda Anco y el (la) asesor(a): Dr. Juan Jesús Soria Quijate

con el propósito de administrar el acto académico de sustentación de la tesis titulado:
"Análisis de sentimientos en Twitter con stacking ensemble para el índice de aprobación política de la presidencia del estado peruano"

del(los) bachiller(es): a) Deivis Ronald García Cercado
 b) William Salvadory Barzola Torres
 c) Juan Manuel Villarroel Lopez

conducente a la obtención del título profesional de:
Ingeniero de Sistemas

El Presidente inició el acto académico de sustentación invitando al (a la) / a (los) (las) candidato(a)s hacer uso del tiempo determinado para su exposición. Concluida la exposición, el Presidente invitó a los demás miembros del jurado a efectuar las preguntas, y aclaraciones pertinentes, las cuales fueron absueltas por al (a la) / a (los) (las) candidato(a)s. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado.

Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Bachiller (a): Deivis Ronald García Cercado

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	16.8	B+	Muy bueno	Sobresaliente

Bachiller (b): William Salvadory Barzola Torres

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	16.8	B+	Muy bueno	Sobresaliente

Bachiller (c): Juan Manuel Villarroel Lopez

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	16.8	B+	Muy bueno	Sobresaliente

(*) Ver parte posterior

Finalmente, el Presidente del jurado invitó al (a la) / a (los) (las) candidato(a)s a ponerse de pie, para recibir la evaluación final y concluir el acto académico de sustentación procediéndose a registrar las firmas respectivas.

 Presidente/a

 Secretario/a

 Asesor/a

 Miembro

 Miembro

 Bachiller (a)

 Bachiller (b)

 Bachiller (c)

ÍNDICE

RESUMEN	5
ABSTRACT	5
1 INTRODUCCIÓN	6
2 METODOLOGIA	7
2.1 Arquitectura del modelo propuesto.....	7
2.2 Materiales y Métodos.....	9
2.2.1 Recopilación de datos.....	10
2.2.2 Preprocesamiento	10
2.2.3 Construcción del método ensamblado.....	10
2.2.4 Evaluación del Método Stackin	16
3. RESULTADOS Y DISCUSIÓN	19
3.1 Resultados descriptivos	19
3.1.1 Características de usuarios de X.....	19
3.1.2 Resultados de Frases Coincidentes con Palabras Clave de Búsqueda.....	20
3.1.3 Análisis de Longitud de Textos por Sentimiento	20
3.1.4 Exploración de las palabras más comunes en el análisis de sentimiento.....	21
3.2 Resultados inferenciales	22
3.2.1 Clasificación mediante los Modelos Base Seleccionados	22
3.2.2 Clasificación mediante el Método Stacking Ensemble	25
3.2.3 Resultados del análisis de sentimiento.....	25
3.2.4 Métricas de Evaluación del Método Stacking Ensemble	26
3.2.5 Comparación del Método Stacking Ensemble y Modelos Individuales.....	30
4. CONCLUSION	30
5. REFERENCIAS	31

Análisis de sentimientos en Twitter con stacking ensemble para el índice de aprobación política de la presidencia del estado peruano

RESUMEN

La creciente complejidad de los problemas políticos de un país destaca la importancia de comprender la percepción pública, para buscar alternativas y estrategias que permitan mejorar el bienestar de la población. Este estudio empleó el método stacking ensemble como una alternativa para mejorar la precisión en las métricas de clasificación y evaluar el índice de aprobación política del estado peruano mediante el análisis de sentimientos en Twitter. Con un enfoque metodológico riguroso que abarcó desde la recolección de datos hasta su presentación, el estudio logró una clasificación eficiente utilizando una muestra de 8724 tuits recolectados de ciudadanos peruanos del año 2023. Se generaron gráficos dinámicos que ilustran las puntuaciones de clasificación de sentimientos, lo que facilita una interpretación intuitiva. La investigación tuvo como objetivo mejorar las métricas de precisión, recall, exactitud y F1-score en la clasificación del análisis de sentimientos. Los resultados de los modelos individuales mostraron que Support Vector Machine alcanzó un (accuracy=75%, MSE = 0.306). Sin embargo, el método stacking ensemble logró una accuracy del 77% y un MSE de 0.2846, lo cual es más óptimo en comparación con los métodos individuales.

Palabras clave: Stacking ensemble, Análisis de sentimientos, Aprendizaje automático, La calificación de aprobación en Twitter.

ABSTRACT

The increasing complexity of a country's political problems highlights the importance of understanding public perception in order to seek alternatives and strategies that can improve the well-being of the population. This study employed the stacking ensemble method as an alternative to enhance the precision in classification metrics and to assess the political approval index of the Peruvian state through sentiment analysis on Twitter. With a rigorous methodological approach that ranged from data collection to its presentation, the study achieved efficient classification using a sample of 8,724 tweets collected from Peruvian citizens in 2023. Dynamic graphs were generated that illustrate the sentiment classification scores, facilitating intuitive interpretation. The research aimed to improve the metrics of precision, recall, accuracy, and F1-score in sentiment analysis classification. The results from individual models showed that the Support Vector Machine achieved an accuracy of 75% and an MSE of 0.306. However, the stacking ensemble method achieved an accuracy of 77% and an MSE of 0.2846, which is more optimal compared to individual methods.

Keywords: Stacking ensemble, Sentiment analysis, Machine learning, The approval rating Twitter