PyCaret AutoML and BiLSTM Accuracy Gaps in Multilingual Emotion Classification

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 9 peer-reviewed papers addressing the following research question: How does the accuracy gap between PyCaret AutoML and BiLSTM for fine-grained emotion classification vary across different language families (e.g., Indo-European, Sino-Tibetan) when applied to. 7 claims were extracted from source literature; 3 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques. Research question: How does the accuracy gap between PyCaret AutoML and BiLSTM for fine-grained emotion classification vary across different language families (e.g., Indo-European, Sino-Tibetan) when applied to non-English datasets?.

2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.8/10.

3 Results

9 papers retrieved. 7 claims extracted; 3 independently verified. Quality review score: 5.8/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
Major social media sites like Twitter, Facebook, and Quora face issues with spam accounts designed to trap users with ma	✓	0.23
The proposed system aims to classify tweets as either 'spam' or 'ham'.	×	0.09
The proposed system aims to evaluate the emotion (sentiment) of tweets.	×	0.09
For spam detection, the study utilizes Decision Tree, Logistic Regression, Multinomial Nave Bayes, Support Vector Machi	✓	0.34
For sentiment analysis, the study utilizes Stochastic Gradient Descent, Support Vector Machine, Logistic Regression, Ran	✓	0.31
For sentiment analysis, the study utilizes deep learning methods including Simple RNN, LSTM, BiLSTM, and 1D CNN models.	×	0.13
The performance of each classifier used in the study was analyzed.	×	0.07

References

- https://doi.org/10.1016/j.heliyon.2023.e20281
- https://doi.org/10.1155/2022/5211949
- https://doi.org/10.63125/edxgjg56