

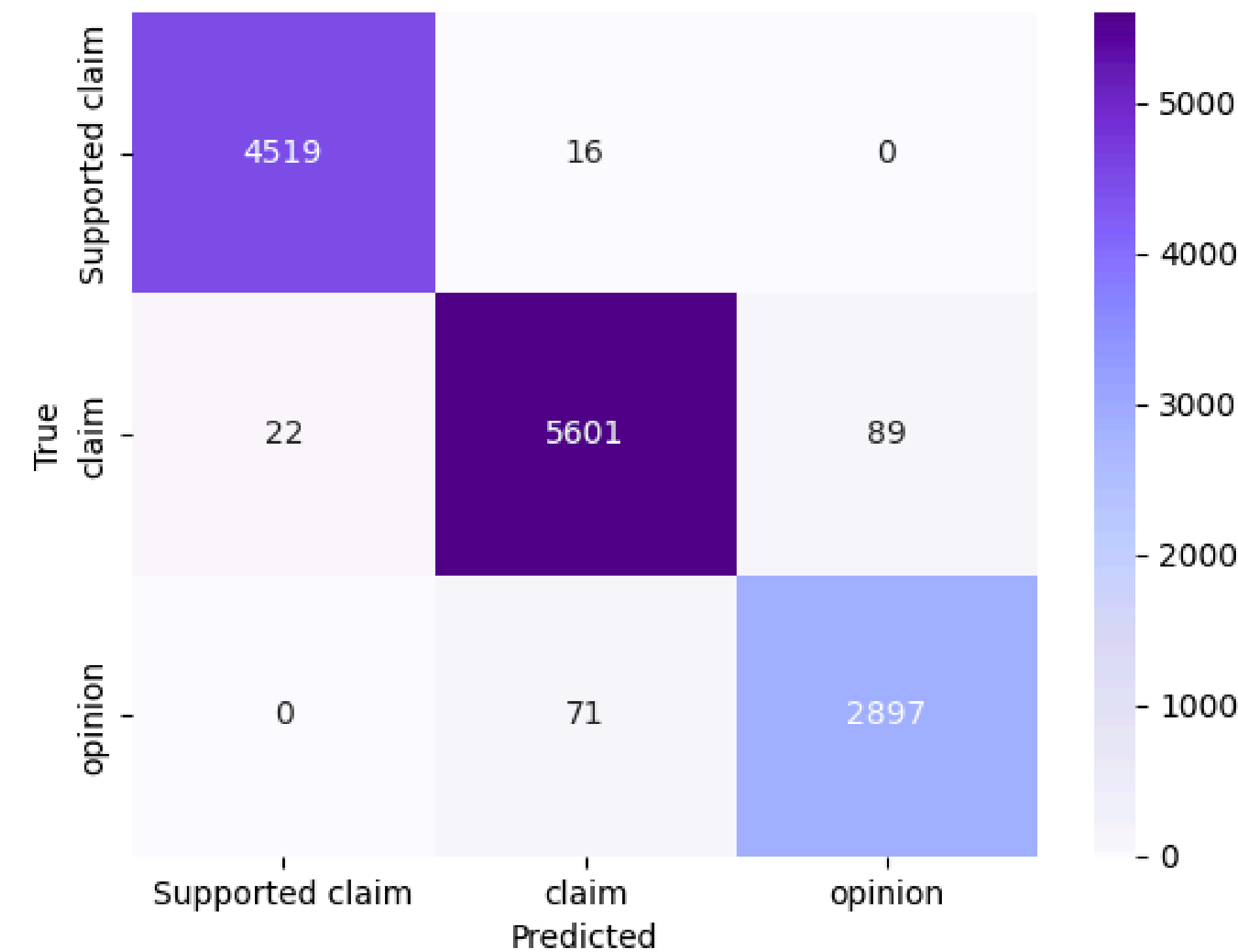


Faktz: An NLP and Machine Learning-Based Platform for Content Credibility on TikTok

Problem Statement

- TikTok's algorithm prioritizes engagement, not accuracy.
- Misinformation on TikTok spreads faster than users can verify it.
- There is no tool that identifies unverified content.

Confusion Matrix - TF-IDF + SVM (Test)



Project Objectives

- Collect, clean, and annotate TikTok transcripts
- Build a complete NLP processing pipeline.
- Train an ML classification model for credibility.
- Deploy the model into a web platform.

Student's Deatils

Mayar Abu Latifa
ID: I25036233

Program: Bachelor's of Information Technology
Email: i25036233@student.newinti.edu.my

System Features

- 01 > TikTok Transcript Processing Pipeline
- 02 > ML-Powered Credibility Classification
- 03 > Source Extraction for Supported Claims
- 04 > Transparent Confidence Score Output

System Strengths

- Trained on 45k+ TikTok + news transcripts
- Handles Gen Z slang & informal language
- Fast pipeline (clean → vectorize → classify)
- High-accuracy SVM classifier
- Transparent confidence scoring

Test Scores

Test Accuracy: 0.985017026106697				
	precision	recall	f1-score	support
Supported claim	1.00	1.00	1.00	4535
claim	0.98	0.98	0.98	5712
opinion	0.97	0.98	0.97	2968
accuracy			0.99	13215
macro avg	0.98	0.98	0.98	13215
weighted avg	0.99	0.99	0.99	13215

Input

"text": "According to a 2023 report from the World Health Organization (WHO), regular physical activity can reduce the risk of heart disease by up to 30%. The American Heart Association also recommends at least 150 minutes of moderate exercise per week, confirming that staying active has long-term health benefits for the heart and overall wellbeing."

Output

```
{  
  "confidence (%)": 98.89,  
  "label": "Supported claim",  
  "resource": "Resources: the World Health Organization (WHO), The American Heart Association"  
}
```

Acknowledgements

Special thanks to my supervisor, Dr. Atif Mahmood, for his continuous support, guidance, and prompt assistance throughout this project.