Predicting Levels of Public Unrest from Social Media Text: A Neural Network-Based Solution

Zen Anthony Pastolero
Arthur John Pagayon
Ryan Cabalfin

BSCS 3-A AI

Instructor: John Christopher Mateo

December 12, 2025

CCS 248 – Artificial Neural Networks

Final Project

Introduction / Problem Description

The objective of this project is to design and train a **Deep Neural Network (DNN)** from scratch to solve a real-world classification problem. Specifically, the project focuses on **classifying public unrest levels from textual data** using emotion-based features derived from user-generated text. This problem is closely related to sentiment and opinion analysis, which is a common application of artificial neural networks.

Artificial Neural Networks are suitable for this task because they can learn complex, non-linear relationships between high-dimensional textual features and categorical output classes.

Dataset Description

The dataset used in this project is derived from the **GoEmotions dataset**, a publicly available dataset containing English text annotated with emotion labels. The dataset was processed and mapped into three unrest-related classes for classification.

Dataset Source:

- GoEmotions Dataset (publicly available research dataset)

Data Type:

- Text data (sentences/comments)

Preprocessing Steps:

- Text cleaning and normalization
- Emotion label mapping to unrest categories
- Conversion of text into numerical features using TF-IDF

Bias and Privacy Considerations:

- The dataset contains publicly available text data with no personally identifiable information.

- No sensitive user information is included.
- Potential bias due to language or cultural context is acknowledged and minimized through balanced class weighting during training.

Neural Network Architecture

The neural network used in this project is a **fully connected feedforward neural network** trained from scratch.

Architecture Overview:

- Input Layer: 5000-dimensional TF-IDF feature vector
- Hidden Layer: Dense layer with 128 neurons
- Activation Function: ReLU
- Dropout Layer: 0.3 dropout rate
- Output Layer: Dense layer with 3 neurons
- Output Activation: Softmax

This architecture allows the model to learn meaningful representations of textual features while reducing overfitting through dropout regularization.

Training Process and Optimizer

The model was trained using supervised learning with backpropagation.

Training Details:

- Optimizer: Adam
- Loss Function: Sparse Categorical Cross-Entropy
- Early Stopping: Enabled to prevent overfitting
- Class Weights: Applied to handle class imbalance

All weights were initialized randomly, and no pretrained models or embeddings were used.

Hyperparameter Tuning

Hyperparameter tuning was performed by training the model using different configurations and comparing validation accuracy.

Configuration	Hidden Units	Dropout	Learning Rate	Validation Accuracy
h128_lr1e-3	128	0.3	0.001	65.37%
h128_lr3e-4	128	0.3	0.0003	**65.39%**
h256_lr1e-3	256	0.4	0.001	65.22%

The configuration **h128_lr3e-4** achieved the highest validation accuracy and was selected as the final model.

Evaluation Results

The final selected model achieved a **validation accuracy of 65.39%**, which exceeds the minimum required accuracy of 50–60%.

Evaluation was performed on a held-out validation dataset using accuracy as the primary performance metric.

Tools and Technologies Used

- Programming Language: Python
- Deep Learning Framework: TensorFlow / Keras
- Feature Extraction: Scikit-learn (TF-IDF)
- Development Environment: Visual Studio Code

Conclusion

This project successfully demonstrates the application of a deep neural network trained from scratch to solve a real-world text classification problem. The model meets all project requirements, including accuracy thresholds, hyperparameter tuning, and compliance with the restriction on pretrained models. The results show that artificial neural networks are effective for sentiment and unrest classification tasks when combined with appropriate feature extraction techniques.