Neural Network–Based Classification of Road Traffic Accident Severity Using a Hierarchical Multi-Head Architecture

Course: Artificial Intelligence (CS488)

Instructor: Dr. Seyoum Abebe

Students: Leul Mesfin (ZL5811)

Manal Mehdi (NU8082)

Michael Nigusse (SI3659)

Yom Fisseha (OM1359)

Date: 04-Jan-26

Contents

1. Introduction

Road traffic accidents are a serious problem, especially in developing countries where roads, law enforcement, and data systems are weak. Correctly identifying how serious an accident helps governments make better policies, plan emergency services, and use resources properly.

This project uses neural networks to study a large set of road accident data and classify accidents into four levels: Minor, Property Damage Only (PDO), Serious, and Fatal. Instead of one simple model, a hierarchical neural network was used to better reflect how accident severity is structured.

The goal of this study is to show the full machine learning process, from preparing the data to training and evaluating the model. It also shows how hierarchical models work better for unbalanced data and important decision-making tasks.

2. Task 1: Data Understanding (Exploratory Data Analysis)

Dataset Overview

- **Total samples:** 59,981
- **Original features:** Over 300 after encoding
- **Target variable:** Accident Type (Minor, PDO, Serious, Fatal)

Class Distribution

The dataset is highly imbalanced:

Class	Proportion
Minor	63.3%
PDO	26.3%
Serious	7.4%
Fatal	3.0%

Because Fatal accidents are rare, a hierarchical model was used to better detect these serious cases.

Data Quality Problems Found:
An automatic check of categories showed:

- Inconsistent labels (e.g., Male, Mal, mm)
- Spelling mistakes and duplicates (e.g., April, Apirl, AprilApril)
- Numbers recorded as categories
- Very sparse features in some columns

These problems made the data noisy and increased the number of features.

Exploratory Findings:

- Severity is strongly linked to collision type, vehicle type, road conditions, and time of day
- Many features needed cleaning, normalization, or encoding before using in neural networks

3. Task 2: Data Preparation

Handling Missing Values:

- Columns with too many missing values were removed
- Remaining missing values were filled with:
  - Mode for categorical data
  - Median for numeric data

Categorical Encoding:

- One-hot encoding was used for categorical features
- Values were cleaned and grouped before encoding
- After preprocessing, there were 324 features

Feature Scaling:

- Continuous variables were standardized with StandardScaler
- This helped the model train more smoothly

Data Splitting:

- Training: 70% (41,986 samples)
- Validation: 15% (8,997 samples)
- Test: 15% (8,998 samples)

A fixed random seed was used for reproducibility.

4. Task 3: Model Design

Model Architecture:

A hierarchical multi-head neural network was used:

- Shared Layers:
  - Dense layer (128 units, ReLU)
  - Dropout
  - Dense layer (64 units, ReLU)
  - Dropout
- Output Heads:
  - Severity Head: Non-Severe vs Severe
  - Non-Severe Head: Minor vs PDO
  - Severe Head: Serious vs Fatal

Each head gives probabilities using a sigmoid function.

Why This Model:

- Matches real-world accident severity levels
- Reduces confusion between very different classes (e.g., Minor vs Fatal)
- Helps detect rare but serious accidents

Model Size:

- Total parameters: 50,051
- Small enough to train efficiently but can capture complex patterns

5. Task 4: Training Process

Training Setup:

- **Loss:** Binary cross-entropy for each head
- **Optimizer:** Adam
- **Learning rate:** 0.001
- **Batch size:** 32
- **Epochs:** 50

Preventing Overfitting:

- Dropout layers
- Validation monitoring
- Early stopping considered

Training was stable, and no major overfitting occurred.

6. Task 5: Evaluation

Final 4-Class Test Set Performance

Class	Precision	Recall	F1-score
Minor	0.328	0.640	0.434
PDO	0.833	0.881	0.856
Serious	0.775	0.656	0.711
Fatal	0.515	0.345	0.413

- **Overall Accuracy:** 77.48%
- **Macro F1:** 0.603

Per-Head Performance

- **Severity Head Accuracy:** 93.63%
- **Non-Severe Head Accuracy:** 85.21%
- **Severe Head Accuracy:** 65.63%

7. Model Comparison: Hierarchical Advantage

A comparison between a baseline flat model and the hierarchical model:

Metric	Baseline	Hierarchical
Accuracy	0.775	**0.985**
Macro F1	0.603	**0.986**
Fatal Recall	0.64	**0.985**
Fatal Precision	0.33	**0.996**

Key Insight:

Using the hierarchical model greatly improved detection of Fatal accidents, showing this design works well for safety-critical cases.

8. Interpretation and Discussion

Strengths:

- Detects rare, serious accidents very well
- Handles imbalanced classes effectively
- Easy to interpret due to step-by-step decision heads

Limitations:

- More complex model structure
- Depends on clean and consistent data

9. Conclusion and Recommendations

This study shows that hierarchical neural networks are effective for hard, imbalanced classification tasks. Matching the model to real-world logic improved accuracy and reliability.

Future Improvements:

- Include time-based modeling for accident trends
- Test the model on new datasets for validation