HiLCoE School of Computer Science and Technology

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

Introduction and Understanding The Problem

Road traffic accidents are a serious issue that causes injuries, deaths and financial damage. Understanding and predicting accident can help authorities improve road safety measures and design better prevention strategies. However it is difficult because it involves many interacting factors. The aim of this project is to predict accident severity using a classification model based on a neural network. There are 4 classification accident types with the target classes being Minor, PDO, Serious, and Fatal accidents. There's strong class imbalance specifically on how rare fatal accidents are, a hierarchical approach is considered the best for learning stability and prioritize the accurate identification of severe cases. Our main goal in this project is to build, train, and evaluate a neural network model. Also showing good data preprocessing, model design, and evaluation practices.

Task 1: Data Understanding (Exploratory Data Analysis)

The dataset given to us includes around 60,000 records with 54 raw features. These features cover driver details, vehicle information, road characteristics, environmental conditions, and accident outcomes. The target variable, accident type, is heavily populated with non severe cases. PDO and Minor cases make up the majority class. Fatal accidents are less than 3% of the data. This imbalance has a significant impact on model design and the model evaluation. Our data analysis showed many missing values, especially in victim-related and post-accident fields. Other variables were inconsistent, such as different spellings or numeric codes for gender. The feature distributions indicated that many variables are categorical with low to medium cardinality, while numerical features, like age and counts of injuries, needed scaling. These noise in the dataset obligated us careful preprocessing, normalization of target labels. Also a clear separation between pre-accident and post-accident columns/features to prevent data leakage was mandatory.

Task 2: Data Preparation

Data preparation aimed to create a clean and ready for the model dataset while still preventing information/data leakage. Missing numerical values were filled in using the median, while categorical variables were filled using the mode or a "Missing" category. The target variable was normalized to four consistent labels: Minor, PDO, Serious, and Fatal. Categorical variables were encoded using one-hot encoding, leading to 324 input features after preprocessing. Numerical features were standardized using StandardScaler. The dataset was divided into 70% for training, 15% for validation, and 15% for testing, using stratified sampling to maintain the original class distribution. Two experimental configurations were created: one with only pre accident data for realistic prediction and another with pre and post accident data to show us the impact of label leakage and establish maximum performance.

Task 3: Model Design

The model features a hierarchical multi-head neural network architecture with a shared backbone and three binary output heads. The shared layers consist of two fully connected layers with 128 and 64 neurons, both using ReLU activation and dropout regularization. This shared representation enables the model to learn general patterns. Instead of a single four class softmax output, our model predicts severity in stages. The first head classifies accidents as Non-Severe or Severe. The second head differentiates between Minor and PDO for non-severe cases, while the third head differs Serious from Fatal for severe cases. This model reflects the natural hierarchy of accident severity and allows the model to concentrate learning on difficult and rare classes, especially Fatal accidents.

Task 4: Training Process

The model was trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and up to 50 epochs. Binary cross-entropy loss was applied for each output head. To address class imbalance, sample masking ensured that each head was trained only on relevant samples, and fatal cases were given higher loss weights to enhance recall. To minimize overfitting, dropout layers set at 0.3 and early stopping with a patience of 10 epochs were used. Training and validation loss curves demonstrated stable convergence, indicating effective regularization. The final model had about 50,000 trainable parameters, balancing expressive power with computational efficiency.

Task 5: Evaluation

The final model was assessed on the held-out test set using metrics like accuracy, precision, recall, F1-score, and confusion matrices. The overall four-class accuracy was 0.7748, and the macro F1-score was 0.6033. The model performed best for the PDO class and weakest for the Fatal class, which is expected due to the significant class imbalance. Evaluation for each head showed strong performance in the first severity split with 93.6% accuracy, confirming that the model successfully distinguishes between severe and non-severe accidents. The severe-type head achieved a Fatal recall of 0.70 within the severe subset, demonstrating the advantage of hierarchical learning.

Per-Head Performance
Severity Head: Accuracy 0.936
Non-Severe Head: Accuracy 0.852
Severe Head: Accuracy 0.656 (Fatal recall 0.70 within severe subset)

Comparing the pre only and pre+post models revealed near-perfect performance when post-accident features were included, clearly showing the impact of data leakage and validating the realism of the pre-only setup.

Metric	Pre-Only	Pre+Post
Accuracy	0.775	0.985
Macro F1	0.603	0.986
Fatal Recall	0.64	0.985
Fatal Precision	0.33	0.996

Task 6: Interpretation and Discussion

The results indicate that hierarchical modeling is highly effective for imbalanced, safety-critical classification tasks. By breaking down the problem into simpler binary decisions, the model improves learning efficiency and allows for focused optimization of rare but important outcomes, such as fatal accidents. The trade-off observed is lower precision for the Fatal class, which may be acceptable in many safety contexts where recall is more important. Limitations of the model include sensitivity to noisy categorical data and slight preprocessing leakage due to scaler fitting before data splitting. Future enhancements could involve stricter prevention of leakage, threshold tuning for each output head, and improved techniques for categorical normalization or embedding. Overall, the model balances interpretability, performance, and practical deployment needs.

Conclusion

This project successfully used a neural network-based hierarchical classification approach to predict road accident severity. Through careful data preprocessing, thoughtful model design, and focused evaluation, the model achieved strong overall performance while emphasizing critical outcomes like fatal accidents. The comparison between realistic and leakage-prone setups underscored the importance of proper feature selection and evaluation discipline. The hierarchical neural network proved to be more effective than a flat classification approach in cases of severe class imbalance. The project shows a well-structured machine learning workflow, from data understanding to model interpretation, and offers a solid basis for future improvements and real-world application.