

CSE 4/587 Fall 2022

Project Phase #3

Team

Name: Raakhal Rapolu

UBIT Name: raakhalr

Name: Kajol

UBIT Name: kajol

Setup Instructions:

Backend:

- Environment details:
 - Python 3.6.15
 - Ubuntu 22.04
- `cd dic_us_accidents`
- Create virtual environment
 - `python -m venv dic_env`
 - `source dic_env/bin/activate`
- Install the python package requirements in the virtual environment created
 - `pip install -r requirements.txt`
- Migrate the django project
 - `Python manage.py migrate`
- Run the Django API
 - `Python manage.py runserver`

Frontend - React:

- Download the zip file and open it
- **Editor:** vscode
 - **Link:** <https://code.visualstudio.com/>

Mac

- Clone the nvm package to the existing shell
 - **Link:** <https://github.com/nvm-sh/nvm#installation-and-update>
- Run below commands in the terminal
 - `curl -o- https://raw.githubusercontent.com/nvm-sh/nvm/v0.39.2/install.sh | bash`
 - `command -v nvm` (if nothing is listed that means nvm is not yet loaded)
 - `nano ~/.bash_profile` (create bash_profile if this command doesnt run)
 - `touch ~/.bash_profile` and then `nano ~/.bash_profile` after that add this line to the script `source ~/.bashrc`
 - `source ~/.bash_profile`
 - `nvm install 10.16.0`
 - `nvm use 10.16.0`
 - To check the installation
 - `Run node -v & Run npm -v`

Ubuntu:

- Sudo apt-get update
- Sudo apt-get install nodejs npm
- To check the installation
 - npm -version
 - nodejs -version

For Windows

- Install Git bash.
 - **Link:** <https://gitforwindows.org/>
 - Then go to select shell in vs code and select default shell as git bash then bash will be the main terminal
- Install node
 - Go to <https://nodejs.org/en/download/>
- Download the 64-bit window installer and run and install
- Restart vscode
- Run node -v and npm -v to check the versions

After setting up the above instruction to run the app, run below command

npm install
npm start

Directory Structure for the project:

```

├── DIC_Project
│   ├── dic_us_accidents
│   │   ├── manage.py
│   │   └── requirements.txt
│   ├── dic_venv
│   ├── Phase_1
│   │   ├── preprocess.py
│   │   ├── visualisation.py
│   │   └── DIC_Phase_1.pdf
│   ├── phase2
│   │   ├── dt_training.py
│   │   ├── knn_training.py
│   │   ├── lr_training.py
│   │   ├── nn_training.py
│   │   ├── rf_training.py
│   │   └── DIC_Phase_2.pdf
│   ├── resources
│   │   ├── my_best_model.epoch08-loss0.381.hdf5
│   │   ├── my_best_model.epoch08-loss0.38.hdf5
│   │   ├── scaler.pkl
│   │   ├── std_scaler.pkl
│   │   └── std_scaler.save
│   ├── severity
│   │   ├── helpers
│   │   ├── migrations
│   │   ├── __pycache__
│   │   ├── services
│   │   └── pred_services.py
│   └── monsters-rolodex (dic react app)

```

Link to dataset: <https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents>

Link to Visualisations:

- <https://app.powerbi.com/view?r=eyJrJoiZGVmMzEzOTYtNjJlZS00ZGQxLWVhYyYyYjdlZjY1NzY1M2M1IiwidCI6IjY1NDY0YTZhLWY4ZWQtdmBiMS05OWUyLTVmNmI1MGEyMDI1MCIsImMiOiN9>
- <https://app.powerbi.com/view?r=eyJrJoiZmJhNDg5YTUtdmBiMS05OWUyLTVmNmI1MGEyMDI1MCIsImMiOiN9&pageName=ReportSection>

Model Selection Criteria

In phase 2 we have trained our processed data on 5 models

Model	Accuracy
Logistic Regression	0.569
K Nearest neighbors	0.71
Decision Tree	0.79
Random Forest Classifier	0.825
Artificial Neural Network	0.86

During phase 2 we have trained the models on a sample of dataset due to limited availability of computational resources, now we have divided the dataset into 5 sample, and performed training neural network on 5 samples, where we have downscaled the number of features by knowing the feature importance from random forest classifier and we have decided using 11 feature and chose the best model based on the accuracy metric, here we chose Neural network among the 5 models that we trained in phase 2 as this was rightly fitting the data, we were able to achieve an accuracy of 0.92 on the test dataset.

In this phase, in our react app, we have used [Microsoft Power BI](#) for visualization and to display the analysis on the dataset.



Problem Statement Analysis

We got the US Accidents data from Kaggle researched the major causes of accidents and attempted to deduce the severity of each one.

- Living in a nation where everything is on wheels involves some amount of driving. If you had just put on your suit and were going to your first interview, what would you least want to happen? If your family has just started a road trip to a national park you've always wanted to see, what is the one thing you absolutely must avoid doing? The response is Car Accidents/Crashes. We learned more about the causes of incidents on the road and how they have changed over time. Where do accidents happen and what are the main causes of car crashes? Whether it is possible to predict how terrible an accident will be and prevent it before it occurs.
- There are patterns in the hours, places, and weather when the majority of occurrences occurred. The seriousness of each accident may be reasonably predicted with the use of numerous classification machine learning methods.
- The following are the questions that will be addressed and which will add to our issue:
 - **When do accidents most frequently happen?**
Daylight against nighttime, Monday through Friday versus weekends, and Peak time
 - **Where are accidents most likely to happen?**
Location, County, City, Zip Code, and Street Side
 - **What kind of weather has the highest accident rate?**
The most dangerous weather for accidents, the three weather conditions that cause accidents the most, and the weather's severity in each accident.

Recommendations

- The US government should limit the number of vehicles on the road. This can be accomplished by enhancing public transit by putting new modes of transportation, such as bullet trains, the hyperloop, and high-speed trains between key cities in California and Texas, into operation. These states have a higher accident rate due to the greater number of automobiles on the roads.
- The government should also consider the weather conditions while road construction and public transport availability. As places with more rain like Florida has the maximum number of accidents, therefore, the roads should be well constructed and proper placements of traffic controls should be there.
- The government should take new automobiles that can withstand high speed entering the market into account when building roadways.
- According to US Accident insights, Weekends had the maximum number of accidents and a major factor for this can be drunk driving. The legal system needs to update and strengthen state-by-state drunk driving sanctions. Additionally, in order to inculcate

discipline among the populace, the government should see to it that those found breaking traffic laws face harsh legal penalties.

Future Work

A significant portion of research has been devoted to the analysis and forecasting of traffic accidents, a major public safety concern. We were able to determine accident-causing factors. Numerous insights into the place, time, weather, and nearby areas of interest of an accident are discovered from this dataset. We were able to determine the most advantageous month, day, and hour for travel. Additionally, it enables us to foresee where accidents are most likely to occur in each state.

In terms of our future work, We can make this platform accessible to the public and state governments so that it is readily available.

This platform can be modified to meet certain issues facing the state's people and administration. To achieve a more current analysis, live data must be loaded into this application and launched on a website.

The public can anticipate an accident location by using our machine-learning model using the source and destination locations, as well as the date and time of transit. The number of accidents that occur in the US may be decreased with the aid of this kind of prediction model. The prediction model may include a number of neural network-based elements that make use of a range of data properties, including traffic-related events, meteorological data, sites of interest, and time information.

