|  |
| --- |
| **FORM 2**  **THE PATENTS ACT, 1970**  **(39 of 1970)**  **&**  **The Patents Rules, 2003**  **COMPLETE SPECIFICATON**  **(See section 10 and rule 13)** |
| **A METHOD TO DETECT SEVERITY OF ACCIDENT** |
| **APPLICANT -1**  **(a). NAME**: REVA UNIVERSITY  **(c). ADDRESS**: REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 9021190211  Email: info@reva.edu.in  **APPLICANT -2**  **(a). NAME**: Mr. Vamshi Krishna T M  **(b). NATIONALITY**: INDIAN  **(c). ADDRESS**: School of Computer Science and Engineering, REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 9148968223  Email: vamshikrishnatm87@gmail.com  **APPLICANT -3**  **(a). NAME:** Mr. Sathish V  **(b). NATIONALITY**: INDIAN  **(c). ADDRESS:** School of Computer Science and Engineering, REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 7795341261  Email: sathish4id@gmail.com  **APPLICANT -4**  **(a). NAME:** Mr. Muttu Handigund  **(b). NATIONALITY**: INDIAN  **(c). ADDRESS:** School of Computer Science and Engineering, REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 7411225697  Email: muttuvittal818@gmail.com  **APPLICANT -5**  **(a). NAME**: Mr. C Yogesh  **(b). NATIONALITY**: INDIAN  **(c). ADDRESS**: School of Computer Science and Engineering, REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 8123493556  Email: yogeshc221101@gmail.com  **APPLICANT -6**  **(a). NAME**: Dr. T Prasanth  **(b). NATIONALITY**: INDIAN  **(c). ADDRESS**: School of Computer Science and Engineering, REVA University, Bangalore, Karnataka, India, 560064  Phone: +91 8667651621  Email: prasanth.t@reva.edu.in |
| **The following specification particularly describes the invention and the manner in which it is to be performed.** |
| **A METHOD TO DETECT SEVERITY OF ACCIDENT** |
| **1. FIELD OF THE INVENTION:**  The invention is related to Computer Science and Engineering. It belongs to the technical field of machine learning, and particularly is related to accident severity prediction modeling method. |
| **2. BACKGROUND OF THE INVENTION:**  The China patent entitled “CSP-CNN model for predicting severity of traffic accident and modeling method thereof” [CN109034264B] describes a model for predicting the severity of a traffic accident and a modeling method. The CSP-CNN model comprises a model input layer inputs a traffic accident gray level image, which gets convoluted and trained by the model to predict the severity of accident using SoftMax activation function. The invention fully considers the space-time relationship, the combination relationship and the deeper internal relationship among the traffic accident characteristics and predicts the severity of the traffic accident.  The China patenttitled “Unmanned vehicle track planning method and system for reducing accident severity” [CN111750887B], the invention provides an unmanned vehicle track planning method and system for reducing accident severity, comprising the following steps: converting the vehicle and environment information from a Cartesian coordinate system to a Frenet coordinate system; generating a reference path by using a cubic spline curve; sampling along a reference path in a scattering point mode, and fitting through a preset secondary polynomial to generate a track cluster; taking the maximum speed, the maximum acceleration and the maximum curvature as control constraints, and removing the tracks which do not accord with the control constraints; performing collision detection on the reserved tracks conforming to the control constraint, and screening the track with the minimum loss function value from the non-collision track cluster through a preset loss function to serve as an optimal track; in the collision track cluster, the track with the lowest severity is screened as the optimal track through an accident severity prediction model; the invention solves the problem that the severity of the accident cannot be reduced when the conventional path planning method faces the unavoidable accident, and improves the safety of passengers when the accident occurs.  The above patents both, are only able to process accident images and extract the features which are responsible for the occurrence of the accident, but they lack in severity of the accident prediction based on these parameters. Our method is able to use these parameters to predict the severity of accidents and also gives us the visualization of the accident-prone locations. And also, our method is able to detect the features which are directly contributing to the accident severity. |
| **3. SUMMARY OF THE INVENTION:**  A method to detect the severity of the accidents, where the data is collected from user for prediction based on various input parameters. The parameters are converted to numerical format using label encoding and then given as an input to the trained model, on the previous data. The model predicts the severity of the accident based on the pattern recognised by the trained model, which gives the output to the user, what is the severity of accident. This information helps in assisting the emergency situation the accident created and also provides insights into the pattern of occurrence of accident severity. |
| **4. BRIEF DESCRIPTION OF THE DRAWINGS:**  Some embodiments of the present invention are illustrated as an example and are not limited by the figures of the accompanying drawings, in which like references may indicate similar elements in which:  FIG. 1 illustrates the proposed approach of accident severity prediction.  FIG. 2 illustrates the processing of the dataset and model training.  FIG. 3 illustrates the analysis done by the trained model in the background.  FIG. 4 illustrates the feature importance based on random forest algorithm.  FIG. 5 illustrates the accident severity ratio of the data.  FIG. 6 illustrates the number of accidents based on year. |
| **5. DETAILED DESCRIPTION OF THE INVENTION:**  A method to detect the severity of accident based on the features, that responsible directly or indirectly for the cause of accident using a classification algorithm, that is regression model and random forest algorithms.  A design of method for prediction of accident severity based on the inputs taken from the users, based on the parameters such as location, weather conditions, road type, vehicle type and year and other parameters. As shown in the Fig. 1 of the architecture design, the input parameters are quantified based on the trained model, and gives the severity of the accident as an output which comprises of the following options, that is slight, severe and fatal. This method will be helpful in prediction of the severity of an accident, which will be helpful for automated vehicles to avoid the accidents, if possible, else it can reduce the severity based on the situation.  The values taken as input by the webpage is then passed to the algorithm which converts the categorical features to numerical values, which then are fed in to the model. The model predicts the severity of the accident, and this information can be used to prevent the future accidents by improving the road infrastructure or to put up some warnings and also try to at least reduce the severity of the accidents as depicted in fig. 4.  The accident prediction model’s preparation is done by collecting the data from United Kingdom official transportation website. The architecture in the fig. 2 represents the steps involved in preparation of prediction of accident severity model. The data preprocessing is done on the data collected by removing the null values, along with removal of the features that does not have any correlation with the identification of accident severity, found by correlation analysis of the data using Random Forest algorithm.  Several data features are in the categorical format, and they cannot be trained by libraries in scikit-learn. These categorical features are converted to numerical, using an encoding method called as label encoding. When the user inputs, they are also label encoded by this algorithm  The dataset is divided into training and testing sets, which eliminates the overfitting problem, and the data which is used for training the model is then sampled using the random oversampling and random under sampling techniques to avoid the skewing of the accuracy based on the classes size. The sampling prevents the result metrics such as accuracy, recall and precision scores to be skewed towards only a single class containing large number of values.  The data is ready at this point and it is time to train the data using the logistic regression algorithm, along with parameter optimization for best parameter prediction using GridSearchCV algorithm. The best parameters are found, which are now used to train the model and get the results for prediction of accident severity. The model is tested for test and train accuracy scores.  The Logistic regression equation can be obtained from the Linear Regression equation. The mathematical steps to get Logistic Regression equations are given below:  (1)  In Logistic Regression y can be between 0 and 1 only, so for this let's divide the above equation by (1-y):  (2)  But we need range between -[infinity] to +[infinity], then take logarithm of the equation it will become:  (3)  The sigmoid function is used in logistic regression as activation function:  (4)  The essential purpose of predicting the severity of the casualties of the traffic accidents is to provide corresponding medical assistance for the personnel involved in the traffic accidents in time, reduce casualties of the accidents, inform the corresponding emergency decision-making departments in time and avoid causing greater property loss. To this end, we further analyzed the predicted severity of casualties of traffic accidents into three degrees: namely, light traffic accidents, serious traffic accidents and fatal traffic accidents. Since correctness is not the only index for evaluating the prediction capability of the model, and in order to combine the practical application scenarios of the model, we introduce accuracy, recall and F1 Score to analyze the traffic accident test set, wherein the calculation formula of the accuracy is as follows:  (5)  wherein TP (true positive) represents a true positive case, that is, the true category is a positive case, and the prediction category is a negative case; FP (false positive) indicates a false positive case, i.e., the true class is a negative case and the predicted class is a positive case.  The recall ratio is calculated as follows:  (6)  FN (false negative) indicates a false negative case, i.e. the true category is a positive case and the predicted category is a negative case.  The formula for F1 Score is as follows:  (7)  Performance metrics of different models for the same data:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Model** | **Precision** | **Recall** | **F1 Score** | **Accuracy** | | Logistic Regression | 92% | 87% | 89% | 81% | | Random Forest | 86% | 87% | 87% | 77% | | XG Boost | 89% | 73% | 80% | 70% |   Visualization techniques, such as graphs, charts, and heatmaps, are employed to present the results of the analysis in a clear and intuitive manner. Interactive visualization tool, which is tableau is used to find the patterns and locations of accidents on real map by the year. It will be helpful in visualizing the accidents according to the year of occurrence and parameters, which are the causes too.  The fig. 6 represents the number of casualties according to the year of the accidents. The figure shows us that the number of accidents has reduced over the passing of each year which is a good sign, but the numbers have to significantly reduce. Thus, our motive in developing this comprehensive model for prediction of accident severity, can be improved to be compatible for the automated vehicles in future. |
| **6. CLAIM:**  **We Claim:**   1. A Method to Detect Severity of Accident comprising:     Collecting the data from the user on the parameters such as weather information, type of vehicle, year of accident, road type, number of vehicles collided, road conditions, and various other parameters;  The data is then converted into numerical data using label encoding and given as an input to the model, that we have already trained using the dataset from United Kingdom official website of road transportation.  The model is specifically trained with the dataset for maximum accuracy and other performance metrics such as F1 score, recall and precision.  The model then predicts the accident severity based on the parameters and gives us an output as severe, slight or fatal.  This is useful for predicting the parameters which are in a combination responsible for the occurrence of the accident severity.   1. The method of claim 1, wherein supervised machine learning models are trained using labeled datasets of accident data. 2. The method of claim 1, wherein the correlation between the dependent features and independent features is identified, and we can find out the conditions on which the maximum number of accidents occur, that helps in categorizing the features which directly lead to the accident severity and indirectly affects the accident severity. 3. The method of claim 1, we also created a visualization of the data, that helps in analyzing the various patterns and features majorly contributing to the accident severity based on the year of the occurrence of the accident. |
| **7. ABSTRACT:**  The invention is based on the classification algorithm, that is logistic regression, which is used for predicting the severity of a traffic accident based on various input parameters such as weather conditions, road type, vehicle type, location, and other similar parameters and a prediction modeling method thereof.Logistic regression is a fundamental classification technique. It belongs to the group of **linear classifiers** and is somewhat similar to polynomial and [**linear regression**](https://realpython.com/linear-regression-in-python/). Logistic regression is fast and relatively uncomplicated, and it’s convenient for us to interpret the results. Although it is essentially a method for binary classification, it can also be applied to multiclass problems. The input parameters by the user will be label encoded and fed into the trained model, which predicts the accidents severity. The output of the accident severity then can be used to interpret the situation of the accident and also predict the future outcomes. |
| **8. DRAWING:**    **Fig. 1**    **Fig. 2**    **Fig. 3**    **Fig. 4**    **Fig. 5**    **Fig. 6** |
| **Vamshi Krishna T M** |