AI-Driven Exploration And Prediction Of Company Registration Trends With Registar Of Companies(RoC)

Submitted By: Nennuru Lepakshi

Register No: 723921243037

Email ID: nennurulepakshi@gmail.com

Building an AI-driven exploration and prediction project involves several key steps:

Exploratory Data Analysis (EDA): In this phase, you'll analyze and visualize your data to gain insights and understand its characteristics. You'll identify trends, outliers, and relationships within the data. Tools like Python's pandas and matplotlib/seaborn can be helpful for EDA.

The study of initial public offering (IPO) markets, their changing trends, and the stock market has been an essential arena of financial analysis over the years. An IPO refers to the mechanism by which private corporations generate capital by offering their shares to public investors while issuing a new stock [1]. It is considered one of the significant transitions in ownership of shares, as the existing private company can offer their shares in the public market to generate more capital. Additionally, IPO allotment is a quick and easy inflow of capital to finance the various ventures of the firm. Moreover, it improves the company's public image once it enters the global market. Publicly listed companies are bound to attract more investors and stakeholders; therefore, the profit generated with an IPO is shared equally among all of the stakeholders.

The prediction of IPO performance in the stock market comes with a set of challenges, such as the fragility of the stock market, irregularity in data, and external socioeconomic factors affecting the IPO market. Motivated by these challenges, we presented a comparative study of four regression models for predicting IPO performance in a market. The four regression models were the KNN Regressor, Decision Tree Regressor, RF Regressor, and Boost Regressor. We also presented an analysis of IPO data, providing essential inferences that allow a better understanding of IPO trends in the current financial market. For that, two standard datasets were identified and then merged into a single dataset by calculating their correlations. Then, the single dataset was pre-processed by using several data preprocessing steps. Then, critical conceptions were carried out using EDA and data visualization, such as correlations, current gains, and feature importance. Further, the regression models were applied to the standard dataset to predict the IPO performance. Finally, the performance of the proposed architecture was evaluated by using various evaluation metrics, such as the MAE, MSE, RMSE, and accuracy. The results show that the Boost Regressor outperformed the other regression models in terms of accuracy, RMSE, MSE, and MAE. The results show that the maximum accuracy obtained was 91.95% by the Boost Regressor.

The area of IPO performance prediction provides a vast scope for future research. Advanced AI models based on DL, federated learning, and transfer learning can be implemented to obtain better prediction results. More features can be incorporated into IPO datasets to train models, such as sentiments in a market about a particular IPO. Additionally, a comparative study can be performed on the IPO performances of different countries, and the impact of one country's market trends on another country's trends can also be studied.

Feature Engineering: Once you understand the data, you'll need to prepare and engineer features for your predictive model. This may involve handling missing data, encoding categorical variables, scaling features, and creating new features to improve model performance.

The uncertain nature of construction projects entails high risk with a diversity of consequences such as project delay, cost overrun, occupational accidents resulting in permanent disability, and fatalities globally [[1], [2], [3]]. Particularly, safety issues in the construction industry cause not only permanent disability, and heavy loss of casualties for workers, but also significant social and economic costs [4,5]. The risk of workers experiencing occupational accidents in the construction industry is higher than in other sectors in most countries [6]. Therefore, there is an urgent need to reduce the number of construction accidents and their severe consequences to provide workers with safer work-life conditions in construction projects.

Permanent disability or permanent incapacity for work is the condition of an injured worker who loses a part of or the entire earning capacity due to an occupational accident. Accidents leading to fatality, permanent disability, and irreversible health effects are considered to be at the top of injury severity levels, referred to as “catastrophic events” [7]. Over 11% of all disabilities occur as a result of occupational injuries globally [8]. As a result, construction accidents that result in permanent disability have a variety of adverse effects, including productivity loss, recruiting cost, compensation cost, compensation claim, and reputation loss [9]. Besides, construction workers who carry out their job to meet the housing needs of future generations are at high risk of not being able to meet their own needs.

Despite all these facts, the condition of permanently disabled construction workers after the accident is seldom examined in the literature regarding the predictive models. Therefore, this study aims to predict the permanent disability status of construction workers post-accident. Based on the diligent investigation of models and approaches proposed in the literature, it is attempted to fill research gaps in several ways and contribute to the well-being of construction workers. Major research gaps and corresponding contributions of the study are:

•It aims to predict construction injury outcomes for the worker regarding permanent disability in a binary classification by using several machine learning (ML) algorithms. In other words, whether the workers would be incapable of performing their works permanently or not as a result of construction accidents is the primary focus of the investigation. The proposed model can be run by construction firms and/or related institutions after an occupational accident as a practical implication. Literature related to permanent disability heavily relies on qualitative studies. On the other hand, limited quantitative studies use several methods and approaches rather than prediction models based on machine learning algorithms.

•A prediction model is proposed based on tree-based ensemble ML algorithms: Random Forest (RF), Adaptive Boosting (AdaBoost), eXtreme Gradient Boosting (XGBoost), and Extremely Randomized Trees or shortly Extra Trees (ET). Tree-based ensemble ML methods have not been comparatively investigated in the construction safety management literature. Past researches frequently highlighted the high performance of tree-based algorithms in the field [6,13,14]. The approach proposed in this study aims to provide a potential guideline for researchers who investigate means to increase the accuracy of prediction models in this area. It also provides insights for safety professionals and policymakers to improve safety management processes and applications in construction projects.

•This study uses the state-of-the-art optimization algorithm, namely genetic algorithm (GA), for parameter tuning in all the adopted ML models. GA has been successfully applied in predictive models in other domains [15,16]. However, it has limited implementations for parameter tuning of ML methods specific to the construction safety management area. Integrating optimization algorithms for the hyperparameter tuning of ML methods provides a powerful means to enhance the accuracy of predictions. Thus, coupling the tree-based ensemble models with GA could aid in developing more robust safety management practices by increasing the performance of applied models.

•Detailed multi-step feature engineering is presented for pre-processing, which includes data encoding, data scaling, dimension reduction, and data resampling. Such a comprehensive pre-processing approach has rarely been applied in this field. Data pre-processing is a fundamental and preliminary step for developing prediction models, which also forms a basis for predicting a target variable. Therefore, an integrated feature engineering procedure through multi-stage pre-processing is aimed to improve the prediction accuracy of enhanced safety of workers in the construction industry.

Predictive Modelling: After feature engineering, you can build your predictive model. You'll need to choose an appropriate algorithm (e.g., regression, decision trees, neural networks) and train the model on your data. Evaluation and validation are crucial to ensure the model's accuracy and generalization.

Predictive modelling is an artificial intelligence (AI) method that uses statistical analytics and data mining techniques to extract valuable insights from previous data and provide our clients with a look into their customers' future behaviour.

Predictive modelling is a strategy for using probability to outline and anticipate certain events, allowing businesses to anticipate problems and propose appropriate solutions. By discovering hidden information in your historical data assets, our predictive modelling experts help you think proactively.

Predictive modelling solutions employ historic insights and historical performance to discover and anticipate the possibility of future trends and the link between broad factors, whereas descriptive analytics techniques summaries prior experiences and search for a single important component to explain behaviour.

Businesses can optimize all aspects of their operations, from acquisition and onboarding to upsell and retention, by identifying what consumers are most likely to do.

Predictive Modelling and Forecasting Solutions

Customers might fall in and out of love with companies for a variety of reasons. A customer centricity strategy is built on staying ahead of consumer preferences and knowing what drives their behaviour in connection to the brand. Furthermore, predictive modelling may be used internally to track critical company stakeholders, such as predicting staff turnover and improving employee engagement.

Predictive modelling aids in determining what consumers are most likely to do next, allowing organizations to plan and implement practical tactics that meet their growth or sales targets. Our company is able to modify customers plans by predicting behaviour, whether it’s through the introduction of complex email sequences, online experiences, or simply optimizing their operations, to guarantee the business has everything in place when the consumer walks through the door.

Hyperparameter Tuning: Fine-tune your model's hyperparameters to optimize its performance. Techniques like grid search or random search can be employed for this purpose.

Deployment: Once you have a well-performing model, you can deploy it in a production environment to make real-time predictions. Tools like Flask or Django can be used for building APIs.

Monitoring and Maintenance: Continuously monitor the model's performance in a real-world setting and retrain it as needed to keep it accurate over time. This is important for ensuring that your predictions remain reliable.

Documentation: Properly document your project, including data sources, methods used, and model details, to make it understandable and maintainable by others.

Remember, the success of your AI-driven project relies on careful planning, thorough data preparation, and robust model development and maintenance.

Thank You!