

## ABSTRACT

### PREDICTING THE POPULARITY OF INDIE GAMES ON STEAM PLATFORM USING CATBOOST ALGORITHM AND SHAPLEY ADDITIVE EXPLANATIONS (SHAP)

By

Mohammad Teddy Syamkalla 20110013

*With the rapid increase in the trend of indie game releases, developers need to increase the potential popularity of the games they release amidst a competitive environment. Previous research using logistic regression and random forest algorithms on the dataset of indie games from the steam platform using all types of features, has less than optimal performance and does not provide knowledge about what features contribute or affect the popularity of indie games. This research aims to predict the popularity of indie games on the steam platform and find out what features affect their popularity. The method used in this research is the CatBoost Algorithm which can handle categorical and numerical data simultaneously with better performance as a prediction model, as well as the Shapley Additive Explanations (SHAP) method to interpret the contribution and influence of features on the model's prediction results. The evaluation results show that the CatBoost model has an overall accuracy of 81%, with precision of 0.83, recall of 0.77, F1-score of 0.80 indicating a balanced ability of the model to identify differences between popularity classes. This is supported by the AUC value of 0.88 where the curve tends to move close to 90 degrees. The SHAP method reveals some of the most significant features that affect the results of predicting the popularity of indie games, such as the existence of steam trading cards category, RPG genre and the use of mac operating system in indie games will increase popularity. Similarly, the number of achievements provided, and the high price will also increase the popularity of indie games. However, the casual genre in indie games will reduce the popularity of indie games.*

**Keywords :** Prediction, Indie Games Popularity, CatBoost, Shapley Additive Explanations, Steam