

Automatic Prediction of Sales in Video Games using Random Forest Algorithm in Comparison with Support Vector Machine Algorithm to Improve Accuracy

INTRODUCTION

- This study aims to improve video game sales forecasts using Random Forest and SVM algorithms, ensuring more accurate predictions and making more informed decisions regarding game development and marketing strategies in the gaming industry
- The problem in video games sales is lack of accurate sales predictions. This may leads to the risk of overproducing or underproducing games, leading to either excess inventory or missed revenue opportunities
- The solution to this problem is identifying the most effective approach for generating accurate sales forecasts in the gaming industry. Moreover, machine learning algorithms like RF and SVM gives better video games sales predictions
- Random Forest Algorithm is an ensemble learning technique that constructs multiple decision trees during training and outputs the mode of the classes (classification) or mean prediction (regression) of the individual trees
- Support Vector Machine is a supervised learning algorithm used for classification and regression tasks. It works by finding the optimal hyperplane that separates different classes in a dataset while maximizing the margin between the classes

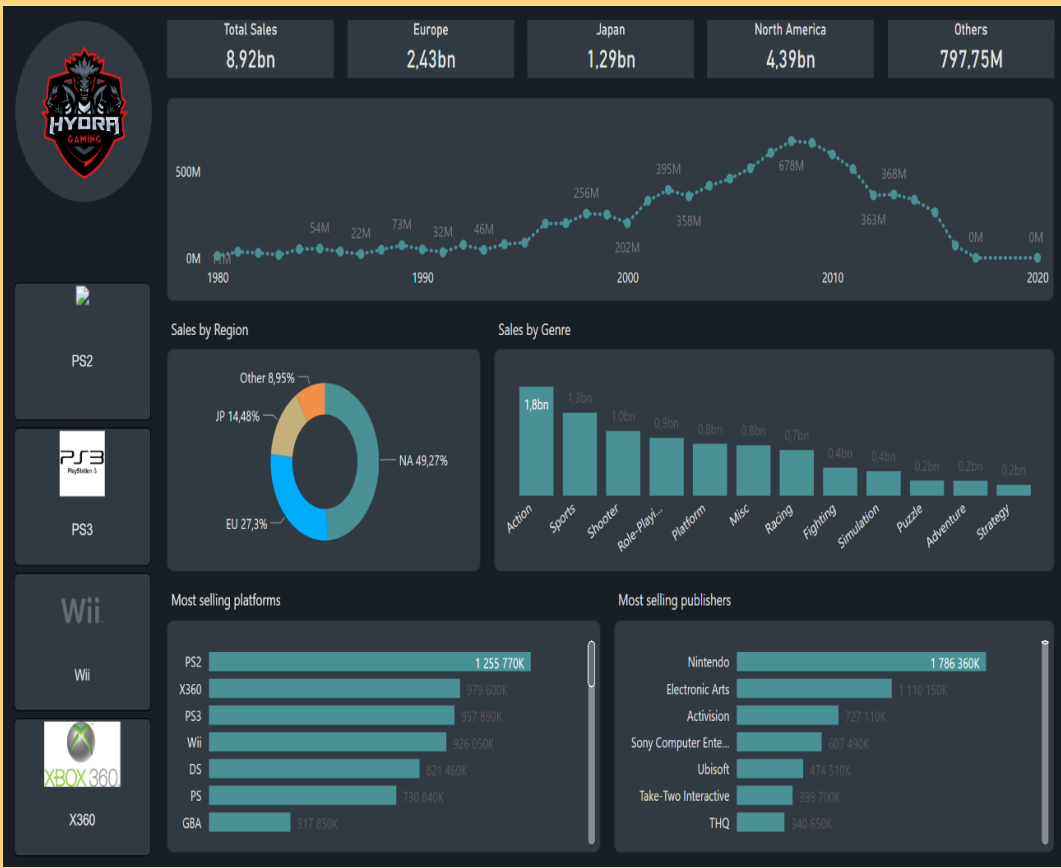


Fig.1 Video games sales prediction

MATERIALS AND METHODS

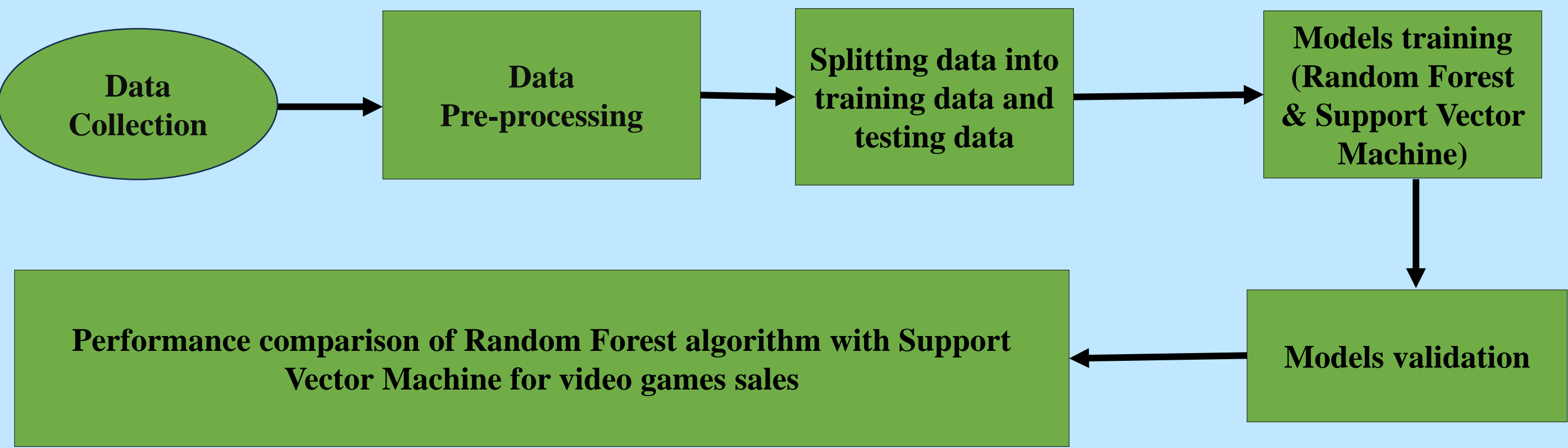


Fig.2 Flow chart of video game sales prediction using RF and SVM algorithms

- Sample Size: 20
- Group-1 (Random Forest): 10 samples
- Group-2 (Support Vector Machine): 10 samples
- The samples are tested in SPSS statistical analysis and outputs have been derived
- The statistical features extracted are count of samples, mean, standard deviation, minimum and maximum
- Non-number data is converted into numerical data for comparison between two algorithms in the dataset

RESULTS

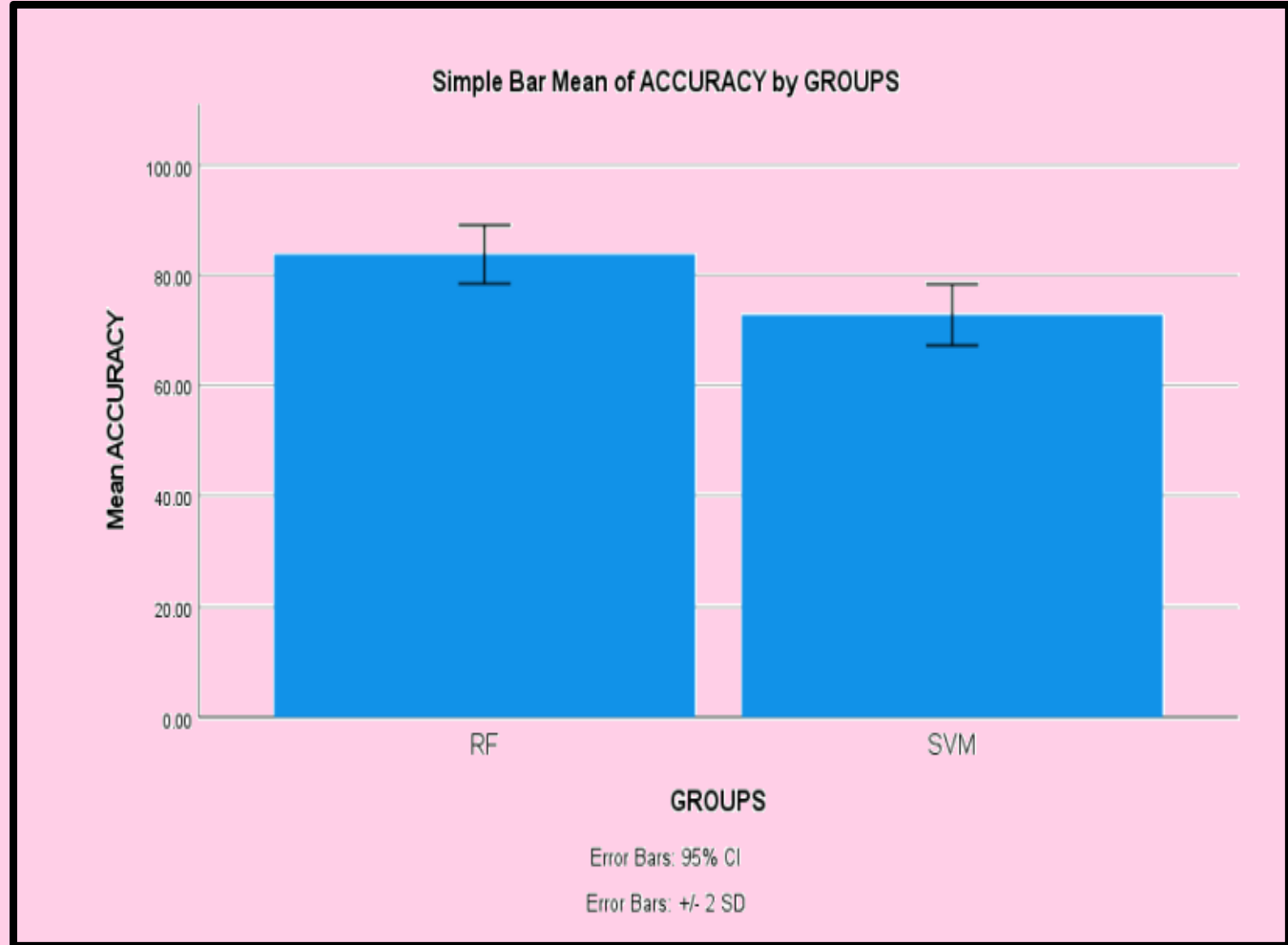


Fig.3 Mean Accuracy Analysis of RF and SVM

Table.1 Group Statistics of RF and SVM Algorithms. It contains mean, standard deviation and standard deviation error

Groups		N	Mean	Standard deviation	Standard deviation error
Accuracy rate	RF	10	86.40	2.658	0.840
	SVM	10	74.49	2.780	0.879

- In the present work, Random Forest Algorithm is compared with SVM algorithm and it depicts that the Random Forest algorithm gives more accuracy
- This Random Forest model attained 86.40% accuracy, on the other hand support Vector Machine Algorithm attained an accuracy of 74.49%

DISCUSSION AND CONCLUSION

- Based on t-test statistical analysis, the significance value of $p=0.001$ (independent sample t - test $p<0.05$) is obtained and shows that there is a statistical significant difference between the RF and SVM. The average accuracy values of two algorithms: Random Forest Algorithm(RF) - 86.40% Support Vector Machine (SVM) - 74.49%
- Both methods contributes to improve accuracy, enabling stakeholders to make more informed decisions regarding game development and marketing strategies
- Moreover, considering the dynamic nature of the gaming industry, ongoing analysis and adaptation of algorithms will be essential to maintain accuracy in sales forecasting
- From the work , it is concluded that the Random Forest algorithm attains the high accuracy when comparing with Support Vector Machine Algorithm in video games sales Prediction
- Moving forward, further research could develop deeper into fine-tuning these algorithms for specific genres or target demographics within the gaming market. Additionally, integrating other advanced machine learning techniques or exploring ensemble methods could offer even greater predictive capabilities

BIBLIOGRAPHY

- Li, Jianbin, Yufan Zheng, Haoran Hu, Junhui Lu, and Choujun Zhan. "Predicting Video Game Sales Based on Machine Learning and Hybrid Feature Selection Method." In 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), pp. 497-502. IEEE, 2021. DOI: 10.1109/ISKE54062.2021.9755343
- Cai, Xiaowei, Javier Cebollada, and Mónica Cortiñas. 2022. “A Grounded Theory Approach to Understanding in-Game Goods Purchase.” PloS One 17 (1): e0262998 DOI: 10.1371/journal.pone.0262998
- Conde-Ripoll, Rafael, Diego Muñoz, Adrián Escudero-Tena, and Javier Courel-Ibáñez. 2024. “Sequential Mapping of Game Patterns in Men and Women Professional Padel Players.” International Journal of Sports Physiology and Performance, February, 1–9. DOI: 10.1123/ijsp.2023-0484
- Etchells, Peter J., Alexandra L. Morgan, and Daniel S. Quintana. 2022. “Loot Box Spending Is Associated with Problem Gambling but Not Mental Wellbeing.” Royal Society Open Science 9 (8): 220111. DOI: 10.1098/rsos.220111
- Doran, John P. 2023. Unity 2023 Mobile Game Development: Discover Practical Techniques and Examples to Create and Deliver Engaging Games for Android and iOS, 2nd Edition. Packt Publishing Ltd. DOI: 10.1007/s10964-024-01950-7