Leveraging Sentiment Analysis of Steam Reviews for Growth in Game Success Metrics

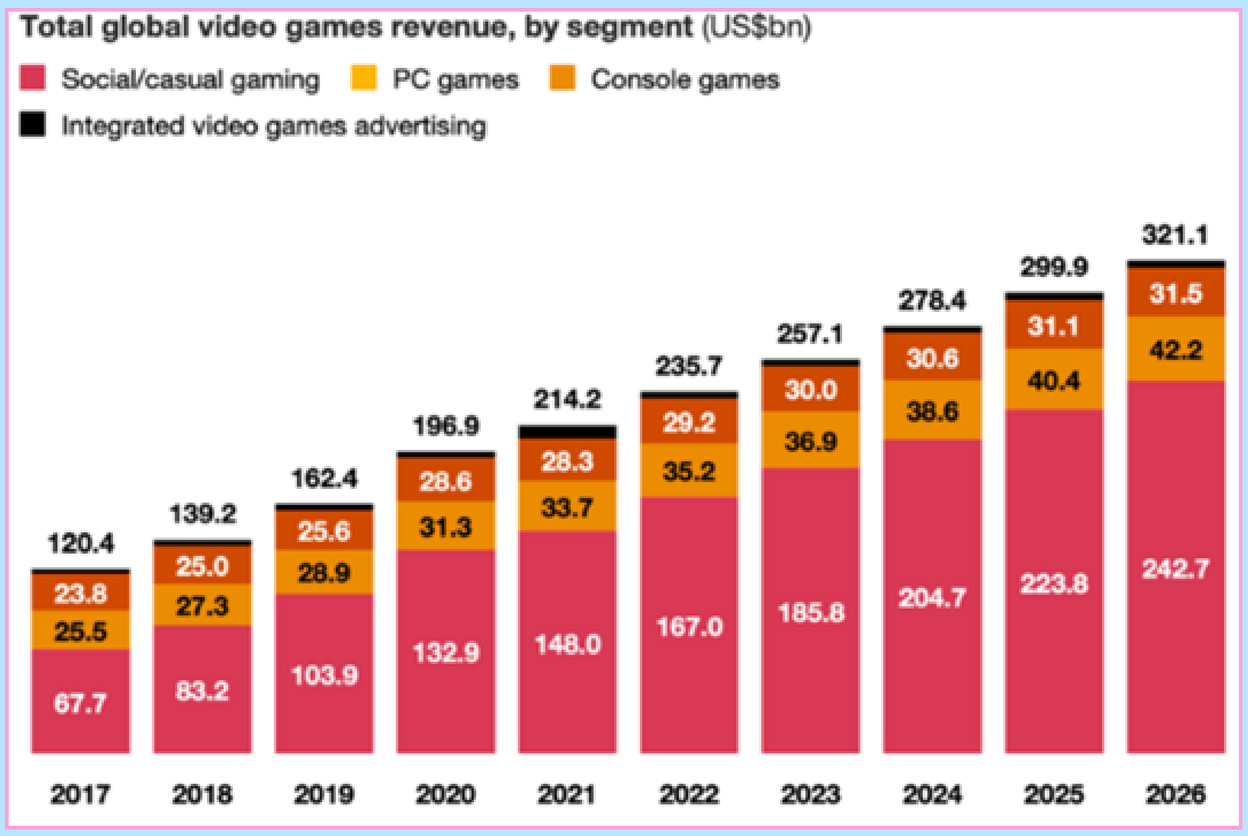
Vedant Tomer

Applied Research Proposal
Master's in Business Analytics
Dublin Business School
Dublin, Ireland
vedanteus@gmail.com

Research Question: Can sentiment analysis of Steam user reviews predict and improve game's success on the Steam platform?

Introduction:

The digital gaming industry is booming, and understanding factors that contribute to a game's success is crucial. This research explores the potential of sentiment analysis in Steam user reviews for improving game success metrics like sales and player retention.



Proposed Methodology:

Data Collection:

- Real-time data from SteamDB, SteamSpy
- Web scraping with Scrapy
- Steam API

Data Preprocessing:

- Text cleaning (remove special characters, digits, punctuation)
- Lowercase conversion
- Stop word removal (NLTK)
- Stemming
- Hyperlink and short review removal

Feature Selection and Vectorization:

- N-gram analysis
- Information Gain (IG) for feature selection
- Term Frequency-Inverse Document Frequency (TF-IDF)

Model Selection and Evaluation:

- Naive Bayes, Support Vector Machines, Decision Tree classifiers
- Grid search for hyperparameter tuning
- Cross-validation for model evaluation

Literature Review:

- **Contextual Understanding via Machine Learning:** Al Mursyidy Fadhlurrahman et al. (2023): Integration of LSTM and CRF in sentiment analysis improves accuracy by retaining more contextual information.
- **Challenges of Online Spam in Reviews:** Pengze Bian et al. (2021): Explores spam detection in game reviews using semi-supervised learning, highlighting the need for robust filtering methods.
- **Historical Techniques and Their Challenges:** Pang et al. (2002): Examines traditional ML algorithms (Naive Bayes, Max Entropy, SVMs) in movie reviews sentiment analysis, noting specific challenges in nuanced language interpretation.
- **Language-Specific Sentiment Analysis:** Thiago Alexandre Salgueiro Pardo (2023): Introduces the SteamBR dataset for Brazilian Portuguese, applying advanced feature extraction for sentiment evaluation.
- **Technical Issue Identification in Reviews:** Mustika Kurnia Mayangsari et al. (2023): Uses KNN, Decision Trees, and Naïve Bayes to predict bug severity from game reviews, aiding in prioritizing developer responses.
- **Determining Review Helpfulness and Engagement:** Zhi Wang et al. (2021): Analyzes predictors of review helpfulness and funniness using Random Forest and Gradient Boosting, enhancing user interaction on platforms.
- **Influence of Gameplay Experience on Reviews:** Dayi Lin et al. (2019): Highlights how playtime influences content and helpfulness of game reviews, offering insights into consumer behavior and marketing strategies.
- **Impact of Game Updates on Player Feedback:** Yang Yu et al. (2021): Uses topic modeling to analyze player feedback on esports game updates, helping developers refine games effectively.
- **Factors Affecting Review Helpfulness:** Lukas Eberhard et al. (2018): Identifies longer reviews and extensive gameplay before reviewing as key factors in perceived helpfulness, guiding potential buyers.

Expected Results:

- Identify the most effective sentiment analysis classifier for game reviews.
- Determine the impact of feature selection techniques on classification accuracy.
- Investigate the influence of preprocessing methods on sentiment analysis.
- Explore the correlation between sentiment and game success metrics.
- Provide actionable insights for game developers to improve design and marketing.

Future Research/Potential Artifact:

- Develop a sentiment analysis tool for real-time game review analysis.
- Explore sentiment analysis for different game genres.
- Investigate the sentiment of in-game communication for game improvement.