

Sentiment Analysis

Game Reviews



Introduction

understand and analyze the reviews provided by gamers on the Steam platform about their gaming experience

Scope

We mainly focused on inspecting nature of every review

Preprocessing

In any sentiment analysis, it is considered to be very important to do data preprocessing before any further analysis

Word Profanity Analysis

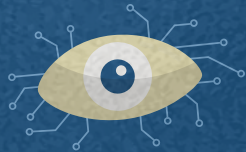
Detecting for any profanity word that could possibly appear in the reviews provided by the gamers

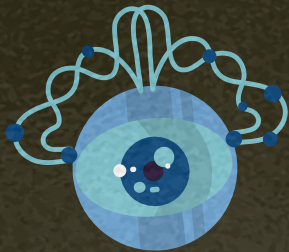
Recsys

Collaborative filtering engine

Conclusion

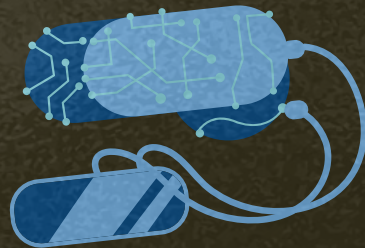
INTRØDUCTION



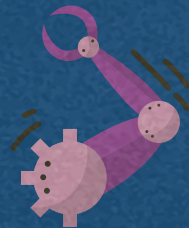
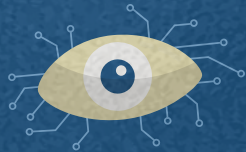


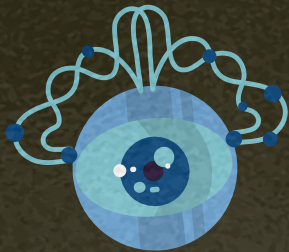
Introduction

In our project, we are trying to understand and analyze the reviews provided by gamers on the Steam platform about their gaming experience



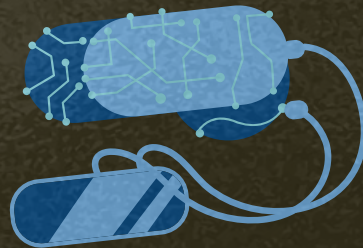
Scope





Scope

inspecting nature of every review





Tools...



Pandas



Numpy



Sklearn



Better Profanity

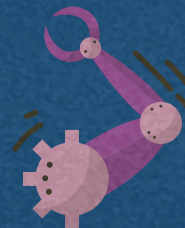


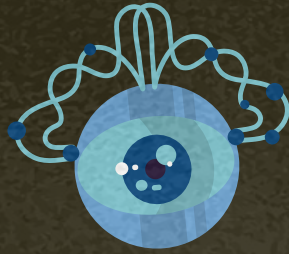
Picke



NLTK

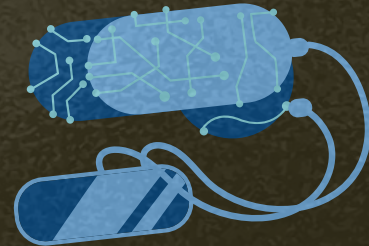
Preprocessing





Preprocessing

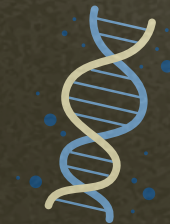
In any sentiment analysis, it is considered to be very important to do data preprocessing before any further analysis



Preprocessing: Distribution of our Data



21million+



56%

Was reviews of **Different** languages

44%

Was reviews of **English** language

Preprocessing: steps...

Cleaning

Remove punctuations,
lowering, numbers etc.



Filtering

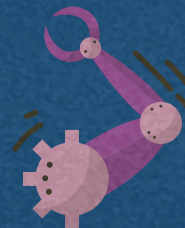
We were only concerned about
the not recommended games

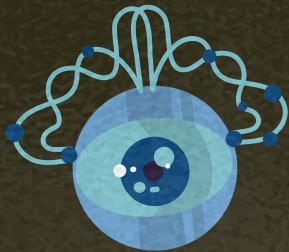
Sampling

Huge data :)



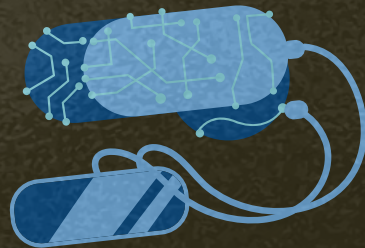
Word Profanity Analysis





Word Profanity Analysis

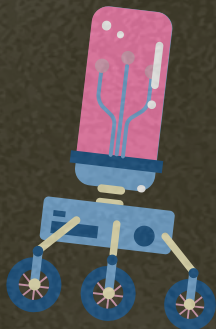
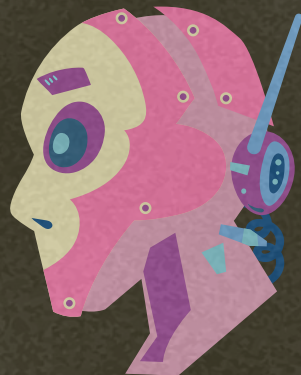
Detecting for any profanity word that could possibly appear in the reviews provided by the gamers



word profanity example...

Review in CS:GO

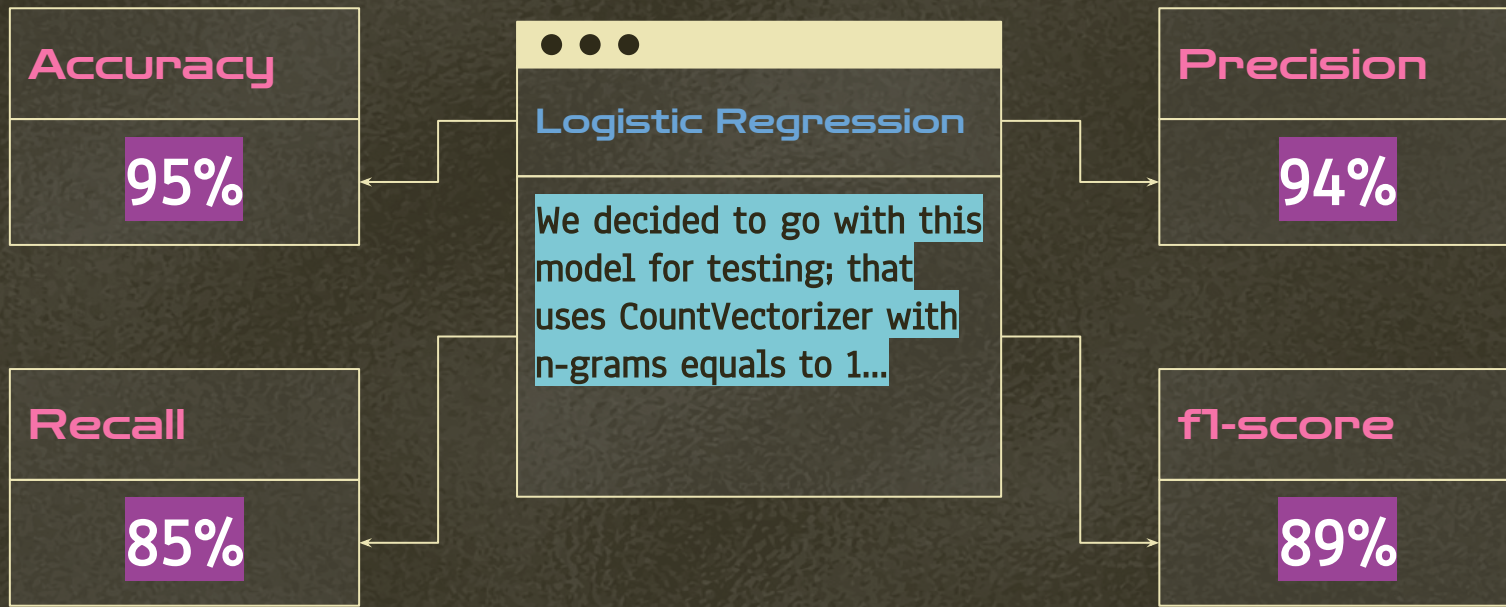
" ..I think the game is a bit lame,
and waste of time :/ "



Review in Dark Souls

" this game is **** ' "

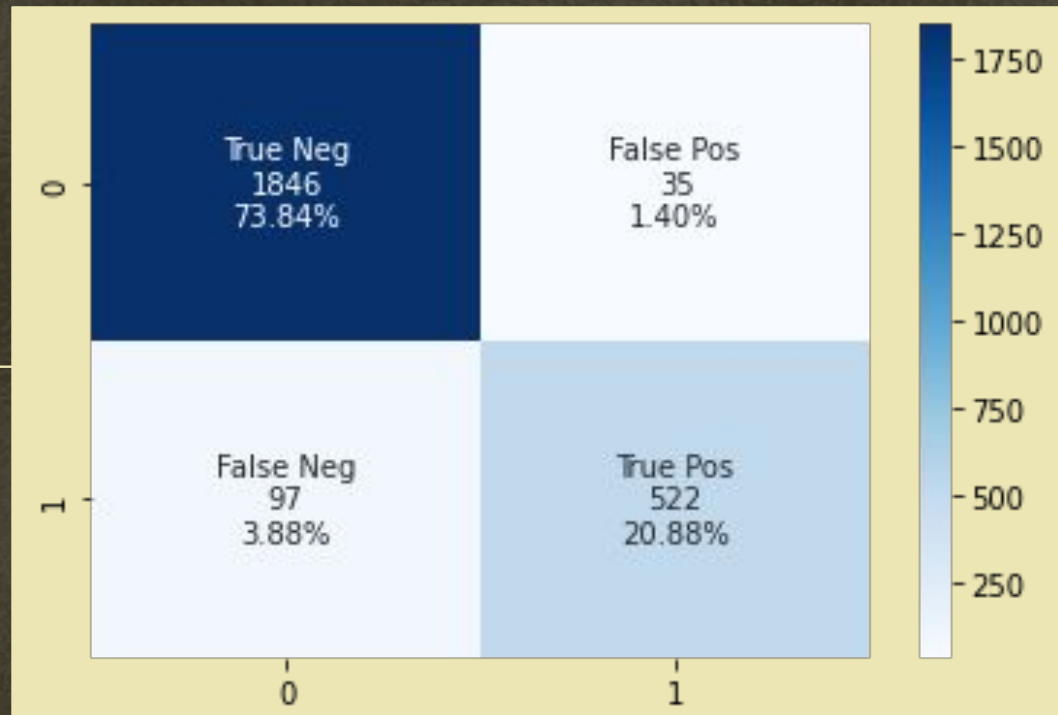
scores...PA:



Confusion Matrix

Profanity Word
1

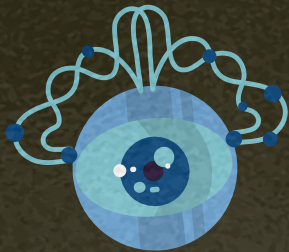
NOT
Profanity Word
0





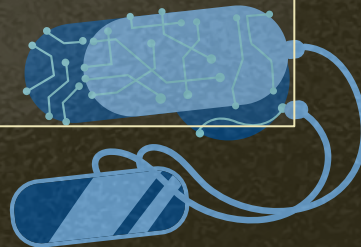
Recsys





Recsys

- Collaborative filtering engine
- Assumed rating
- PCA
- Eigenvalues > 60
- Unexplained!
- Therefore, it is not implemented



Conclusion and lessons learned

1. Large dataset
2. Scope changes
3. Recommendation system problems
 - User-game similarity matrix
 - Different algorithms didn't work except PCA
 - Standardizing ordinal data?
4. Pickling

**Thanks
Questions?**

