# Predicting ratings of Mobile apps

**Parag Verma**
pverma6@ncsu.edu

**Rohan Patil**
rrpatil3@ncsu.edu

**Sayali Dhavale**
sdhaval@ncsu.edu

## 1 Project Idea

The idea of this project is to use Regression models to predict the rating of a given app on Google Playstore using the reviews given by its users and other features like the size of the app, its price, how many people downloaded the app. The reviews are in english and we are planning to use natural language processing to extract meaningful features from the reviews and to feed those features to our regression models.

We are using Linear Regression as our baseline model and would progressively try to build better and better non-linear models for this task. We will also try to compare two models, one which uses sentiment polarity (given in the dataset) and the other which uses features extracted from the reviews using natural language processing. For the task of NLP would use pretrained GloVe or Word2Vec embeddings and a simple LSTM encoding scheme for extracting meaningful features from the reviews. Finally we will compare our final models with our baseline model.

## 2 Dataset

We are using the 'Google Play Store Apps' dataset from Kaggle, which has been generated by scraping the Google Play Store site. `https://www.kaggle.com/lava18/google-play-store-apps/home`

## 3 Software Used

We will be using Python(3) to implement the model. These are some of the Python packages and modules that we might use in our project: Pandas, NumPy, NLTK, Scikit-learn, Keras

## 4 Work Division

We are working on this project in a group of three. We have decided to divide our work as follows. Two people will work on implementing the actual models and one person will work on the NLP preprocessing and feature extraction. All of us will work on literature survey.

## 5 Midterm Milestone

We are hoping to complete the preprocessing and evaluation of the baseline model. We will start building the general framework for other models as well.

## References

[1] D. Monett and H. Stolte, "Predicting Star Ratings based on Annotated Reviews of Mobile Apps," Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, 2016.

[2] D. Monett and H. Stolte, "Predicting Star Ratings based on Annotated Reviews of Mobile Apps," Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, 2016.

[3] H. Malika and Elhadi M.Shakshuki, "Mining Collective Opinions for Comparison of Mobile Apps", The 11th International Conference on Future Networks and Communications (FNC 2016)