

Contents	
About Company	

THEORY

01 Recommender Systems.....	1
1.1 Clustering algorithms.....	2
1.2 Challenges with recommender systems.....	3
02 Machine Learning.....	5
2.1 Machine Learning: Algorithms Types.....	5
2.2 Supervised Learning.....	6
2.3 Unsupervised Learning.....	7
2.4 Semi-Supervised Learning.....	8
03 Text Analytics and Data Analytics Lifecycle	
3.1 Introduction.....	9
3.2 Text Analysis Steps.....	9
3.3 Lemmatization, and Stemming.....	10
3.4 Data Analytics Lifecycle Diagram.....	12

IMPLEMENTATION

Product – VOTEGIRI.IN	13
04 Tools and Libraries Used.....	14
4.1 Python and Pandas.....	14
4.2 Jupyter Notebook.....	15
4.3 Scipy.....	15
4.4 Scikit-learn.....	16
4.5 Natural Language Toolkit.....	17
4.6 Mysql connector.....	17
05 Building Recommendation system for VOTEGIRI.IN... ..	18
5.1 Calculating Tf- Idf	19
5.2 Removing less important words.....	21
5.3 Stemming words.....	22
5.4 Clustering.....	24
5.4.1 Kmeans Clustering.....	25
Result	29
Conclusion	31
Appendices	
References	
Internship Evaluation Report	