NLP Projects

ID	Title	Description	Dataset
1	Sentiment Analysis of movie reviews	Design a model to apply Sentiment Analysis of movie reviews using sentiment polarity datasets (positive and negative). The overall process of model training and testing should be recorded and displayed using a proper chart. The classification results for the best model should be displayed too.	http://www.cs.cornell.edu/people/pabo/movie-review-data/
2	NewsGroup classification	Build a model to classifying news data into various categories through text classification. The overall process of model training and testing should be recorded and displayed using a proper chart. The classification results for the best model should be displayed too.	http://qwone.com/~jason/20Newsgroups/
3	Sentence to Sentence semantic similarity	Can you identify question pairs that have the same intent or meaning? Dataset: Quora question pairs with similar questions marked	https://www.kaggle.com/c/quora-question-pairs/data
4	Fake News Detection Project	Fake news spreads like a wildfire and this is a big issue in this era. You can learn how to distinguish fake news from a real one. You can use supervised learning to implement a model like this.	https://drive.google.com/file/d/1er9NJTLUA3qnRuyhfzuN0XUsoIC4a-_q/view
5	Topic Modeling	Topic modeling is a method for finding a group of words (i.e. topics) from a collection of documents that best represents the information in the collection of text documents. It can also be thought of as a form of text mining - a way to obtain recurring patterns of words in textual data. The topics identified are crucial data points in helping the business figure out where to put their efforts in improving their product or services.	https://drive.google.com/file/d/1VaZrCOrfh2qURMd5_nJ04b4fD5SE3vhx/view?usp=sharing
6	movie recommender	predict the user movie rating	https://grouplens.org/datasets/movielens/