# ReadMe

This project a Senior Capstone of Maniz Shrestha, an undergraduate majoring in Computer Science at Earlham College. The Capstone titled “Detecting Fake News using Sentiment Analysis and Network Metadata” builds a machine learning model to detect fake news. The project contains a web application that implements the fake news model.

This repository has two main directories – TrainModel and FakeNews. TrainModel contains the implementation of the machine learning model and FakeNews contains the Flask application. The following is a brief description of important files in this repository:

* FakeNews
  + /static directory contains static web files.
  + /templates directory contains the template for the front-end.
  + init.py – python package initialization.
  + config.py – configuration setup for database in Flask.
  + customTransformer.py – module to implement Machine Learning Pipeline.
  + scrape.py – module to scrape news content and metadata (such as Facebook analytics and domain rank).
  + retrainModel.py – module to retrain the machine learning model from new data from the database.
  + models.py – Flask Models implementation in MVC framework.
  + views.py - Flask Views implementation.
  + docModel – exported document to vector model object.
  + rfclf.pkl – pickle file that contains the trained classifier.
* TrainModel
  + customTransformers.py – a module to implement classes for machine learning pipeline.
  + Single\_test\_pipeline.py – code to train the model and evaluate the performance with FakeNewsCorpus dataset.
  + Eval\_2datasets\_pipeline.py – code to train the model on FakeNewsCorpus and evaluate the model with both the FakeNewsCorpus and Getting Real about Fake News Dataset.
  + add\_metadata\_to\_dataset.py – code to acquire data from Facebook and Open PageRank and export two dataset with metadata (FakeNewsCorpus and GFRN).
  + create\_Dataset.py – code to clear, filter and randomly select FakeNewsCorpus dataset.
* manage.py – this is the file that is used to run the flask webserver.

To run the webserver, type **python manage.py runserver** in the terminal in the main directory.