Shivam Singh
Data Scientist

Data Scientist with proven expertise in Machine Learning and Deep Learning with expertise in building end-to-end machine learning solutions. I aim to coordinate my knowledge, hard work and self-motivation gained in the corporate world to seek new challenges that would accelerate my skills and abilities.

✉ shivamcse17818@gmail.com   📱 7007320632   📍 Basti, India

EDUCATION

Bachelors in Computer Science
Government Enginnering College

07/2017 - 08/2021   *Sonbhadra , India*

Courses
- B.Tech , 7.6 SGPA

WORK EXPERIENCE

Data Scientist
3 Analytics Pvt Ltd

06/2020 - 02/2021
Social media monitoring for Adverse Events

- Business Objective: Monitor social media, to check for adverse events caused post vaccination for covid or for any drug in focus. In addition identify genuine tweets, with a set of pre-defined conditions
- Extracted and perform text preprocessing to clean data on social media data from twitter.
- Defined annotation approach; perform Cluster analysis and topic modeling on the text from tweets, to understand the existing pattern and aid in labeling
- Checked the comparative performance of multiple transformers - BioBERT_Disease_Ner, ClinicalBERT_Ner, and define road ahead to identify Symptoms in the text.
- Based on the labeled data, train a Bio-Bert model for classification of the tweets as the ones with adverse event and non-adverse events. Achieved an f1-score of 66.65%.
- And After Training the model created a web interface using ReactJS and deployed this project on AWS production server.
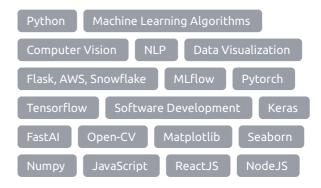
Data Scientist
Hobasa India Pvt Ltd

03/2021 - Present
Email Classification

- Business Objective: Monitor Outlook Inbox, to categories mail within specified category and move into dedicated folder after classification.
- Extracted and perform text preprocessing to data on outlook inbox.
- Defined annotation approach; perform Cluster analysis and topic modeling on the text from inbox, to understand the existing pattern and aid in labeling.
- Checked the comparative performance of multiple Models - BERT(transformers), LSTM, XGboost in the text.
- By help of outlook graph API, created folders based on labels and moved mails to the dedicated folders.

SKILLS

Python   Machine Learning Algorithms   Computer Vision   NLP   Data Visualization   Flask, AWS, Snowflake   MLflow   Pytorch   Tensorflow   Software Development   Keras   FastAI   Open-CV   Matplotlib   Seaborn   Numpy   JavaScript   ReactJS   NodeJS

PERSONAL PROJECTS

Cassava Leaf Disease Classification
- Our task is to classify each cassava image into four disease categories or a fifth category indicating a healthy leaf in this competition.
- I have trained 5 models(EfficientNet b-4,b-5,b-3 and Resent 50d, 200d) using Bi-Tempered loss function and use heavy augmentation.
- I ranked 74 out of 3900 in this competition.

Mechanisms of Action (MoA) Prediction
- Our task is to improve the algorithm that classifies drugs based on their biological activity.
- I have combined simple NN, tabnet, and 2 head-resnet after performing PCA, some feature engineering, and feature selection methods.
- I ranked 94 out of 4373 in this competition.

ACHIEVEMENTS

Kaggle Competition Expert (12/2012 - Present)
Ranked among the top 1% (830th out of 168000) on the leaderboard of the world largest data science community - Kaggle(https://www.kaggle.com/shivam17818)

LANGUAGES

English
Professional Working Proficiency

Hindi
Native or Bilingual Proficiency

INTERESTS

Volleyball   Online Games