RAVINDER KUMAR TANWAR

DATA SCIENTIST | KAGGLE 2X EXPERT

✉ rtanwar616@gmail.com
🌐 https://www.kaggle.com/ravijoe
📞 +918847095358
📍 Mohali,Punjab
in https://www.linkedin.com/in/ravi-tanwar-12bb3811a/
🐙 https://github.com/ravijoe/

Skills

PROGRAMMING

Python
Flask
PyTorch
Numpy
SciPy
Matplotlib
Seaborn
Scikit-Learn
Docker
API
Data Structures

AUTOMATION / WEB SCRAPING

Selenium
Scrapy
BeautifulSoup

STATISTICS AND PROBABILITY

Statistics
Probability Basics
Random Variables
Probability Distributions
Hypothesis Testing
Modelling

PRODUCT METRICS

A/B Testing

MACHINE LEARNING/DEEP LEARNING

K Nearest Neighbors
Random Forest
Linear and Logistic Regression
NLP

CLOUD COMPUTING

AWS

DATA VISUALIZATION

Tableau
PowerBI
Plotly

Education

PEC University Of Technology   May 2015 to Apr. 2019
Bachelor Of Engineering - Computer Science

Springboard Data Science   Aug. 2019 to Feb.
Career Track   2020
Completed a 6 month Data Science Career Track
that includes 650+ hours of hands-on curriculum,
with 1:1 industry mentor oversight, and
completion of 2 capstone projects. Mastering
skills in Python, SQL, Data Analysis and
Visualization, Hypothesis Testing, and Machine
Learning.
https://www.linkedin.com/in/ravi-tanwar-
12bb3811a/detail/overlay-
view/urn:li:fsd_profileTreasuryMedia:
(ACoAAB22TbMBRg2Zo_J4IhcvP6ogMdec7Vd2bFo,1593288157466)/

Employment

SearchUnify   Panchkula
Data Scientist   Nov. 2019 to Current

- Started as a part of Analytics team . Worked in Data Fetching(using APIs) , Pre-Processing(using Python) and Reporting(using Tableau and Microsoft Power BI)
- Mediums used include **Google Analytics and Microsoft Dynamics365**.
- **Currently leading NLP department of the team.**
- **Was responsible for mentoring two juniors who completed POCs on Anomaly Detection and Next Sentence Prediction within 2 month timeframe**
- Responsible for **hiring** new Data Scientists for the team .
- Worked on **Unsupervised Text clustering methods, Embedding techniques like Tfidf,Word2Vec,LSTSMs,Transformers,BERT,Longformers etc.**
- Worked on Recommendation System for recommending similar queries from a **dataset size >1 million rows** and **reduced the memory used by previous algorithm by 2 folds**
- Currently working on **Anomaly Detection and Imbalanced Classification** for detecting Escalated user tickets with **Model giving 90+ Recall along with proper Business understanding.**

Code Garage Tech   Mohali
Python Developer   May 2019 to July 2019

- Worked on Web Scraping of Dynamic Websites using **Scrapy , BeautifulSoup , Selenium in Python** .Some websites included **one of the largest places for wine collection and one of the largest State-wise school database for India** .
- Used python for automating tasks which included but not limited to **quick conversion of word files into pdfs ,automated liking and deleting of posts on Facebook.**
- Used OpenCV for Object And Image Detection Algorithms For example Harr Cascade.

Council of Scientific and Industrial Research-Central Scientific Instruments Organisation   Chandigarh
Intern   Jan. 2018 to June 2018

- Worked in the Biomedical Instrumentation unit at CSIR-CSIO Chandigarh and
- Took part in working toward Development Of Rehabilitation Tools using Unity3D. Positive feedback from doctors at GMCH ,Sector 32 was received

Projects

Song Popularity Prediction   Aug. 2019 to Oct. 2019
- Data-set was picked from **Kaggle** , which contained 250k songs collected using Spotify API.
- Features included were **acousticness , danceability,** loudness ,energy etc.
- Target variable was popularity between 0-100
- Model used was Random Forest .
- **Shared the results with YouTube musicians and got positive feedback**

Movie Recommendation System   Sept. 2019 to Jan. 2020
- Used TMDB dataset from Kaggle containing 5000 movies .
- Demographic and Content based Techniques were used.
- Cosine Similarity was used for finding similar movies .
- Successfully deployed the model on Python Flask Framework
- **Posted the demo video on Linkedin and got 200+ likes and 10000+ views**

News and Stats + Dashboarding using Twilio and Tableau
- Made an alert system for latest covid-19 news and stats .
- Used **BeautifulSoup** to scrap data from https://www.worldometers.info/coronavirus/.
- **Following the completion of data fetching the stats and news are sent to user's Whatsapp number using Twilio API.**
- **Used the built product for self use and notifications ; got decent results**

Web App for identifying Fake News using NLP
- Made a classification system for identifying Fake News among a given list of news .
- After model training and testing ;deployed the same model using Flask Web Framework in Python.
- **User can input a given text as News and get prediction telling whether news is fake or not .**

Speech Sentiment Analysis
- Worked on a Voice Recognition system that provides the Sentiment for the speech.
- Used **PyAudio** module to convert text to speech.
- Converted Speech was then fed into **VaderSentiment** Analysis package which is in NLTK library.
- **Suggested the idea to present company and got approval** for the same since Voice based methods methods improve security.

Awards

Datacamp · Data Scientist with Python Track   Dec. 2019

Hackerrank4 Hackerrank Skill Certificates
https://www.hackerrank.com/certificates/01b077a48963

Coursera · Introduction to Data Science in Python - University Of Michigan

Activities

Wrote article on 'Grammar Checking using BERT' which got featured by Analytics India Magazine
https://www.linkedin.com/posts/ravi-tanwar-12bb3811a_how-to-use-bert-transformer-for-grammar-checking-activity-6704088266787971072-Dnmu

Made 10k+ followers on Linkedin by Posting Data Science related content
https://www.linkedin.com/mynetwork/invite-connect/connections/

Achieved trending status for 3 posts related to #datascience

Collaboration with Ken Jee and Andrew Mao
Collaboration with Ken Jee(ranked as one of the top Data Scientists in the World) and Andrew Mao(ranked as one of the top Data Scientists in the World) for their YouTube channels https://www.youtube.com/channel/UC23emuGbNM7twofQIrEgPBQ