Aman Sawarn

https://www.aman-sawarn.github.io
sawarn69@gmail.com | +91 8076927879 |

SKILLS

PROGRAMMING/TOOLS
Python and iPython Notebook
Tensorflow
SQL
HTML
Google Cloud Platform (GCP)
Git and Github
MongoDB
Shell Scripting (Bash)
Plotting, cleaning and data Mining tools

DEEP LEARNING
Neural Networks
Natural Language Processing
Long Short Term Memories (LSTM)
Convolution Neural Networks (CNN)
Attention/Transformers
BERT

MACHINE LEARNING
Logistic Regression
Decision Trees
Random Forests
XGBoost and other Ensembles
Recommendation System
Natural Language Processing
Collaborative Filtering
Content Filtering

EDUCATION

DEEP LEARNING
Coursera- deeplearning.ai
| Stanford University

MACHINE LEARNING
Coursera
| University of Michigan

B.TECH
Maharaja Agrasen Institute of Technology
June 2020 | New Delhi, IN
B.Tech in Electrical and Electronics Engineering

LINKS

Github:// **amansawarn**
LinkedIn:// **amansawarn**
Twitter:// **@aman_sawarn**

EXPERIENCE

HT MEDIA LTD. | Data Scientist
June 2020 – | Gurgaon, IN

- CV Parser
  - Built an in-house CV Parser for to replace to third party CV Parser for **shine[dot]com-A HT Media company** .
  - Cleaned, explored and trained models on large datasets and built **SaaS product** for in-house and market use.
  - Saved **0.01 $** per job apply for the recruiting firm, **contributing three percent of total cost-cutting during pandemic**.
  - **Patent** filed for the product by company, with me as one of the inventors
- Recommender System
  - Built a Job recommender system for **shine[dot]com**.
  - Performed A/B testing for different algorithms, hypothesis and business metrics for new updates on portal.
  - **Increased Average Time interaction with product by 11 percent** and **increased Job applies by 6 percent**
- Fraud Detection
  - Implementing **Recruiter Fraud Detection model** for Shine .
  - Identifying the current drawbacks in the on-going verification cycles to ensure optimum risk mitigation.
  - Analysing tiers of risk and dispute processes for the bad actors on the product platform and implementation for metrics optimization.

PROJECTS

REDDIT FLAIR TAG PREDICTION

- Scrapped Data from **Reddit India** website
- Performed data cleaning, exploratory data analysis, stemming, lemmatization to make data ready for modelling.
- Build models using Logistic Regression, SVMs, Decision Trees and Random Forests, and did model analysis using confusion matrix, f1 scores and AUC scores.
- Model deployed link **here**. and documentation **here**

PUBLICATIONS

MEDICAL REPORT GENERATION USING X-RAY IMAGES

- Worked at the intersection of Computer Vision and Natural Language Processing to **generate medical reports from chest X-Ray images**, to help in **medical risk profiling and diagnosis**.
- Worked on multi-stage architectures and attention mechanism to increase sensitivity and specificity towards information gathered in X-Ray images and generate better reports.
- Imroved existing BLEU-1, BLEU-4 and ROUGE scores using least number of parameters. Paper currently under review.