

Raman Shinde

<https://github.com/Raman-Raje>

Email : raman.shinde15@gmail.com

Mobile : +91 9595161238

CAREER OBJECTIVE

Results-oriented professional with 4+ years of experience in python development and expertise in Machine-learning, Deep Learning, Computer Vision and Natural Language Processing (NLP).

CAREER SUMMARY

- Currently pursuing my career as a Data Scientist with core expertise in Data analysis, visualization, building Machine Learning based solutions.
- My work has revolved architecting ML-driven solutions in products/platforms and getting them live into production.
- Expertise in various ML algorithms(KNN,K-means,Naive Bayes,Logistic Regression,SVM,DT,Random Forest,GBDT etc.) and scikit-learn.
- Experience and expertise in building Deep Learning models using Neural Networks such as LSTM/RNN, CNN, Transformer, Bert, Autoencoders, Memory Networks, Seq2Seq etc.
- Experience in implementing research papers and model building using TensorFlow, Keras, Pytorch for model building. Flask for APIs for model serving.
- Have been using Numpy, Pandas, Matplotlib, openCV, SciPy for data cleaning, processing and visualization.

EXPERIENCE

- **Xpanxion** Pune, India
Data Scientist Jan. 2020 - Present
 - **Digital Access Hub(Python/DL/ML):**
- **Siemens R&D** Pune, India
Product Development Engineer Dec. 2018 - Present
 - **Automation Designer(Python/C++):** Working as a part of sequence designer team. Working on NX as a product developer for Siemens.
- **TCS** Pune, India
Software Developer Dec 2015 - Nov 2018
 - **Application Developer(Python):** Developed an application for client NCRA for Monitoring and Controlling of Giant Metrewave Radio Telescope. Coding for I/O operation. Monitoring H/W devices. Debugging the issue and modifying the GUI as per requirements
 - **Production Management:** Worked in QAPM for client Morgan Stanley. In QAPM team, I had worked in ED&S (Enterprise data and services) team which deals with different applications like CRD, SRD, TS, and PM

PROJECTS

- **Deep Learning**
 - **Neural Machine Translation using Attention mechanism (NLP):** Task is to implement Machine Translator. Attention was used to deal with longer sequences. Data cleaning and output labels were padded with start and end tokens before feeding to n/w.
 - **Neural Question Answering(NLP + Attention + Machine Reading Comprehension):** Objective is to find correct answer for given question and context pair. Implemented Stanford Attentive Reader. SQUAD v1 dataset was used for this project. Various binary and NLP features were used to get the best results. Compared the final results with fine tuned BERT model.
 - **Image Captioning with Flickr 8k dataset. (CNN+RNN+Transfer Learning):** Task is to generate caption for an image. Pre-trained Inception n/w on imagenet dataset was used in combination with RNN to generate caption for the given image. One-to-many RNN is used to train the n/w. For each image 5 standard caption are provided. The data was converted into sequence before feeding the n/w.
 - **Human Activity Recognition:** Task is to detect the human activity based on data from smartphone sensor. The dataset was obtained from kaggle. The problem was solved using both **Machine Learning as well as Deep Learning**. Got **96%** accuracy using machine learning model with already given feature set. We have used LSTM model to solve the same problem using sequential raw data from sensors and got accuracy of **91.82%**.

• Machine Learning

- **Netflix Movie Recommendation System (Collaborative based recommendation):** Objective was for the given movie and user predict the rating given by him/her to the movie. The dataset was obtained from kaggle. Matrix factorization was used to get similarity matrices. Tried and tested various ML models to get minimum Root Mean Square.
- **Movie Plot Synopses with Tags(Topic Modelling):** Objective of the case study was to predict the tags for given movie plot synopsis. Dataset was obtained from Kaggle. Various machine learning models were tried and tested with OvR classifier to get the best results. Logistic regression with Topic modelling gave best accuracy best micro f1 score.
- **Stack Overflow Tag Prediction:** Objective is to predict as many as tags possible with high Precision and Recall. The dataset was obtained from kaggle. The given problem is **multi-label classification problem**. The dataset contains features such as Id, Title, Body and Tags. Data preprocessing and cleaning was done to remove html tags and hyperlinks. Micro-Averaged F1-Score was used as performance metric as mentioned on Kaggle.
- **Amazon fashion discovery engine (Content Based recommendation):** The objective is to recommend similar apparel products in ecommerce. The dataset was obtained with help of amazon product advertising API. 183k data points was obtained from product category womens top. Similarity is used to recommend products. Tried and tested various combination of weighted vectors to get the best results.

CERTIFICATIONS/INTERNSHIP

- Applied Machine Learning course at Applied AI. (Jan 2018 to May 2019)
- Completed Stanford Statistical Learning (Self-Paced) course.
- Completed Deep Learning Specialization course from Coursera
- Internship at IARE, Aurangabad on Industrial automation. (May 2014 - Jun 2014)

TECHNICAL SKILLS

- **Languages:** Python, C++, C
- **Data Analysis:** Pandas, Numpy, Matplotlib, Seaborn, openCV
- **ML /DL Toolkit:** Keras, scikit-learn , tensorflow, pytorch

EDUCATION

- **B.Tech** in Electronics and Telecommunication from SGGSIE&T, Nanded with **CGPA 7.7** (2011 - 2015)
- Class XII (HSC), form Maharashtra State Board of Education with **83.33%** (2009 - 2011)
- Class X (SSC), form Maharashtra State Board of Education with **90.92%** (2008 - 2009)