

Arunkumar Venkataramanan

#14/6, Pothaiya Naidu Street, Kaveripattinam, Krishnagiri, Tamil Nadu – 635112
arunkumar.ramanan@gmail.com | +91 90080 41191 | arunkumarramanan.github.io
Kaggle: arunkumarramanan | LinkedIn: arunkumarramanan | GitHub: ArunkumarRamanan

SUMMARY

- I'm a **Machine Learning Practitioner** (also known as **ML Software Engineer** and **Data Scientist**) has over **6 years** of experience in the fields of **Machine Learning, Deep Learning, Artificial Intelligence, Data Science, and Predictive Analytics & Modeling**.
- I founded and managed **DeepBrainz**, an Independent AI Team Project as "an **AI Lead**" who's been solving various business and real-world problems across industries using **Cutting-edge Technologies** and the **State of the art AI** in **Computer Vision** and **NLP**.

EXPERIENCE

Machine Learning Practitioner (ML Engineer & Data Scientist-ML)

(July 2013 – Present)

DeepBrainz (An Independent AI Team Project)

Bengaluru, India

- Advanced the **SOTA** in **generative images creation** based on **MiFID** as **50.42142 CV** by training **GAN** models such as **BigGAN, StyleGAN, DCGAN, ACGAN, RaLS BigGAN** on **Stanford Dogs Dataset** along its Annotations. (TF/PyTorch)
- Automated **multi-label audio tagging** that scored **0.759 CV** as per **label-weighted label ranking AP** by building, training custom **CNN** models on **Freesound & Yahoo FlickrCC** datasets with augmentations to preprocessed **Mel-Spectrogram**. (PyTorch)
- Predicted demand for an online advertisement** as **RMSE 0.216 CV** by designing, implementing **LGB, Ridge** models, **MLPs, RNNs**, on **Avito's Ad** datasets with pre-trained **word vectors, DenseNet, vgg19** for feature extraction. (sklearn XGB TF Keras)
- Classified and Forecasted the future web traffic** for approximately **145,000** **Wikipedia** pages using **RNN seq2seq, LSTM/GRU** with attention by **Time Series Analysis and Inference** that evaluated on **SMAPE** as **test score 35.48065**. (TF Keras)
- Recommended events to the users** on **Event Recommendation Engine** datasets from **Kaggle** as **0.712 CV** over mean **Average Precision mAP @200** by building, training **Logistic Regression, RF, k-Means** models for feature transformation. (sklearn)

PROJECTS

Large Scale Computer Vision and NLP with Deep Learning

(Jan 2019 – Oct 2019)

- Diagnosed diabetic retinopathy** as **0.936129 test score Quadratic Weighted Kappa** by developing, fine-tuning, augmenting, ensembling **EfficientNet, SE-ResNet, SE-ResNeXt/50/101, SE-Net154** on large set of retina images. (PyTorch)
- Detected Objects Automatically** as mean **AP 0.59 CV** by using pre-trained **ResNet/EfficientNet, NAS-FPN & RetinaNet** models on large-scale **Open Images v5** with **TFX** pipeline for training, testing, deploying models on **TPUs**. (TF)
- Recognized Images** as **Global AP @ k 0.2515 CV** by fine-tuning and augmenting pre-trained **SE-ResNeXt50 & Faster R-CNN** models on large-scale **Google Landmark & CVDF** datasets. (PyTorch)
- Detected toxicity and minimized Bias in Toxicity Classification** as **test score 0.947** of custom **AUC** metric by developing custom **LSTMs** with pre-trained **BERT, XLNet, GPT2** models on **Jigsaw's dataset** and preprocessed embeddings. (PyTorch)
- Improved gender-fairness for gender bias pronoun resolution** by designing **BERT** based models on **GAP Coreference Dataset** with preprocessed embeddings by augmentation measured as **multi-class log loss of 0.1916 test score**. (TF Keras)

EDUCATION

Independent Coursework - MOOCs

(July 2013 - Present)

Coursera	Udacity	edX & Others	Stanford University	MIT	Others
<i>Stanford University:</i> <ul style="list-style-type: none">AlgorithmsMachine Learning <i>deeplearning.ai:</i> <ul style="list-style-type: none">Deep LearningTF in PracticeUToronto Learn to Program SeriesUCSD Big Data	<ul style="list-style-type: none">Machine Learning EngineerNanoDegree (ND)Data Scientist NDData Engineer NDDesign of computer programs, NorvigSoftware Testing & Debugging	<ul style="list-style-type: none">UCSDx Data ScienceIBM Python DSColumbiaX Business Analytics <i>CS50x HarvardX:</i> <ul style="list-style-type: none">Computer ScienceWeb ProgrammingCMU Distributed Systems	<ul style="list-style-type: none">Machine LearningMining Massive DatasetsCNNs for Visual RecognitionNLP with Deep LearningInformation Retrieval	<ul style="list-style-type: none">Design and Analysis of AlgorithmsComputational thinking & Data ScienceMITx: Stats & Data ScienceDeep Learning	<ul style="list-style-type: none">Google AI - Education: Learn with Google AIDatacamp Track: Data Scientist with PythonKaggle Courses: Faster Data science education

B.Tech. Information Technology

University College of Engineering Tindivanam

(Aug 2009 - July 2013)

Anna University, Chennai, India

SKILLS

- Programming:** Expert in **Python, Java, SQL**; Have Experience with **C++, C, C#, Shell script, JavaScript, R, MATLAB**
- Libraries & Frameworks:** Expert in **TensorFlow & Ecosystem Keras PyTorch scikit-learn numpy pandas scipy** | **Python Matplotlib MXNet XGBoost LightGBM**; Experienced in **MapReduce Spark Hadoop ecosystem OpenCV NLTK SpaCy genism Caffe2 fastai Plotly ONNX**; Familiar with **Beam Flume Sqoop Hume Pig HBase Hive Kafka Storm Knime Weka RapidMiner Neo4j caret**
- Tools:** **Colab Anaconda PyCharm VS Code Excel Linux Git MySQL Tableau MongoDB Cassandra RStudio GCP AWS Azure**

AWARDS & HONORS

- Top **0.03%** ranked **38th/102k** & **49th Kernels Master** and Top **0.2%** ranked **235th/122k Competitions Expert** on **Kaggle**
- Won **Bronze Medals** on **Kaggle: 62nd/468** in **Inclusive Images Challenge**, **74th/201** in **Open Images 2019 – Visual Relationship**, **90th/521** in **iMet Collection 2019 FGVC6**