

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**) have over **6 years** of experience in the fields of **Machine Learning, Deep Learning, Artificial Intelligence, Data Science, and Predictive Analytics & Modeling**.
- I founded and manage **DeepBrainz**, an AI Initiative as **AI Chief and Managing Director** who's been solving various business & real-world problems across industries using **Cutting-edge Technologies** and the **State of the art AI** in **Computer Vision** and **NLP**.

EXPERIENCE

Founder & Machine Learning Practitioner (ML SW Engineer & Data Scientist)

(May 2015 – Present)

DeepBrainz Technologies Private Limited

Bengaluru, India

- Identified the answers to real user questions about Wikipedia page content based on Google Natural Questions dataset that is evaluated on **micro F1** as **0.71 CV** using **BERT** embeddings and **LSTM/GRU** models. (TF 2.0)
- Advanced the **SOTA** in **generative images creation** based on **MIFID** as **50.42142 test score** 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)

Independent ML Practitioner & Consultant

(July 2013 – March 2015)

- 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 **35.48065 test score**. (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** over **Quadratic Weighted Kappa** by fine-tuning, augmenting, ensembling pre-trained **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)
- Detected **toxicity** and minimized **Bias in Toxicity Classification** as **0.947CV** 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
Stanford University: <ul style="list-style-type: none">AlgorithmsMachine Learning deeplearning.ai: <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 CS50x HarvardX: <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 Python & RKaggle Courses: Faster Data science education

B.Tech. Information Technology, University College of Engineering Tindivanam, Anna University, Chennai (Aug 2009 - July 2013)

SKILLS

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

AWARDS & HONORS

- Top **0.03%** ranked **38th/104k** & **49th Kernels Master** and Top **0.2%** ranked **235th/124k** Competitions **Expert** on **Kaggle**.
- Won **Medals** on **Kaggle**: **62nd/468** in **Google AI Inclusive Images Challenge**, **74th/201** in **Google AI Open Images 2019 - Visual Relationship**, **90th/521** in **iMet Collection 2019 FGVC6**.