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 **ML Practitioner** (also as **ML Software Engineer** and **Data Scientist**) have over **6 years** of experience in AI/ML/DL/DS fields.
- I founded and manage **DeepBrainz**, A **DPIIT Recognized AI Startup** as **Managing Director** & **AI Chief** who's been solving various business & real-world problems across industries by **Cutting-edge Technologies** & **State of the art AI** in **Computer Vision, NLP**.

EXPERIENCE

Founder & **Machine Learning Practitioner (Data Scientist** & **ML SW Engineer)**   (Jan 2015 – Present)
DeepBrainz Technologies Private Limited , A DPIIT Recognized AI Startup Company   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** score as **0.81 CV** using pre-trained **BERT** and **LSTM/GRU** models. **(TF 2.0)**
- **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 **TPU**s. **(TF)**
- **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, **MLP**s, **RNN**s, 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)**

Independent ML Practitioner & **Consultant**   (Aug 2013 – Nov 2014)

- **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 – Nov 2019)

- **Advanced** the *SOTA* in **3D Object Detection** for **Autonomous systems** and **Generative Images Creation** as follows: **(TF/PyTorch)**
  - **Detected 3D objects** on Lyft's **nuScenes** /**KITTI** as **0.353 CV mAP @ IoU** by **Point Cloud** models such as **PointRCNN, VoxelNet**
  - **Created Generative Images** as **49.4214 CV MiFID** by **GAN** models such as **BigGAN, StyleGAN** on Stanford Dogs Dataset.
- **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 toxicity** and **minimized Bias in Toxicity Classification** as **0.947CV** of custom **AUC** metric by developing custom **LSTM**s 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: • Algorithms • Machine Learning *deeplearning.ai:* • Deep Learning • TF in Practice • *UToronto* Learn to Program *Series* • *UCSD* Big Data	• Machine Learning Engineer NanoDegree (ND) • Data Scientist ND • Data Engineer ND • Design of computer programs, *Norvig* • Software Testing & Debugging	• *UCSDx* Data Science • *IBM* Python DS • *ColumbiaX* Business Analytics *CS50x HarvardX:* • Computer Science • Web Programming • CMU Distributed Systems	• Machine Learning • Mining Massive Datasets • CNNs for Visual Recognition • NLP with Deep Learning • Information Retrieval	• Design and Analysis of Algorithms • Computational thinking & Data Science • *MITx*: Stats & Data Science • Deep Learning	• *Google AI - Education*: Learn with Google AI • *Datacamp* Track: Data Scientist with Python & R • *Kaggle* 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/106k** & **49th** *Kernels Master* and Top **0.2%** ranked **231th/126k** 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*.