Sanidhya Mangal

615-955-8605 mangalsanidhya19@gmail.com /sanidhyamangal /sanidhyamangal

Education

Vanderbilt University *(Nashville,TN)*, MS in Computer Science; *3.94/4.0* **December 2022**

Medi-Caps University *(Indore,India)*, B.Tech in Computer Science; *8.41/10.0* **May 2020**

Technical Skills

Development Tools: Python (Pandas, Scikit-Learn, Numpy, Tensorflow, Keras, Pytorch, Django, Flask), Bash, Scala
Analysis Tools: SQL (Presto, MySQL, Oracle, DynamoDB, ORM), MS Excel, MS PowerBI, Tableau, Hadoop
Deployment Tools: Docker, Kubernetes, GIT, AirFlow, AWS (Lambda, EC2, EKS, ECS, RDS, S3, Sagemaker)

Experience

Asurion **May 2022 – August 2022**

Data Science Intern *Nashville, TN*

- Developed a real-time feature to influence expert behavior during upsell in a call with a 6% improvement in sales.
- Assembled DS life-cycle: ideation, opportunity sizing**(SQL)**, modelling, deployment**(AWS)** & exposure in A/B testing.
- Analyzed business processes and engaged with stakeholders to translate requirements into analytics problems.
- Trained adversarial Roberta-BERT **(PyTorch)** to induce robustness in text-classification with small dataset.
- Devised out of the box metrics to measure, compare & analyze **(Python)** performance of different ML models.

Maize Zhou Lab, a biomedical/healthcare-based lab working on deep learning **July 2021 – May 2022**

Research Assistant *Nashville, TN*

- Contributed to research projects by developing CNN toolkit **(TensorFlow)** to facilitate genome filtering on long and short reads, leading to multiple publications.
- Translated research problem into end-to-end MLOps pipeline, including data processing, modeling, and evaluation.

Engineerbabu **June 2020 – June 2021**

Machine Learning Engineer *Indore, India*

- Engineered a CNN **(TensorFlow)** based system to perform prognosis of lung and colon cancer with 0.92 AUC.
- Supervised a team of six that reduced inference time by 30% for machine learning models by improving the ML pipeline.
- Designed and deployed **(Docker)** web framework **(Django)** in production for edge AI operations & report generation.
- Optimized decision-making process by 12% through EDA, including ETL, hypothesis testing, and statistical analysis.

Greater Kailash Hospitals, healthcare division leveraging Data Science **January 2020 – April 2020**

Machine Learning Engineer Intern *Indore, India*

- Advised on selecting the best modeling techniques to diagnose lung X-ray images for Pneumonia with 94% precision.
- Deployed **(EC2)** web app that provided a second opinion to physicians and reduced approx. 6% of false negatives.
- Improved baseline performance by 20% using transfer-learning based approach to fine-tune MobileNetV2 **(Tensorflow)**.

Projects

Self-SupervisedGAN | *PyTorch, GAN, Self-Supervised Learning, Computer Vision, Generative Modelling*

- Generated high-fidelity images and improved VanillaGAN's performance using self-supervised pre-training.
- Added self-supervision to prevent forgetful discriminator which aids in better convergence and prevents mode collapse.

Interpretable-Bert | *PyTorch, Transformers, Interpretability, Named Entity Recognition, Classification, NLP*

- Designed probes to leverage pre-trained BERT representations to perform named entity recognition on the input text.
- Analyzed contextual representations to examine how pre-training task affects the linguistic knowledge in transformers.

Semi-Supervised Domain Adaptation | *Domain Adaptation, Computer Vision, Unsupervised Learning*

- Explored a research project on how different pre-training methods affects image classification in domain adaptation.
- Augmented different representation learning methods described in paper "Surprisingly Simple Domain Adaption".

AutoMLify | *Machine Learning, Python, JAX, Gaussian Process, Linear Regression, KNN, SVM, Decision Tree, XGBoost*

- Implemented library for autotuning ML algos: classifiers & regressors with bagging & boosting for structured data.
- Extended the library to support neural networks and optimized training time by 10% through parallelization.

Publications

- Yunfei Hu, **Sanidhya Mangal**, Lu Zhang, and Xin Zhou. "Automated filtering of genome-wide large deletions through an ensemble deep learning framework." Methods (2022).
- **Sanidhya Mangal**, Aanchal Chaurasia, and Ayush Khajanchi. "Convolution neural networks for diagnosing colon and lung cancer histopathological images." arXiv preprint arXiv:2009.03878 (2020).
- **Sanidhya Mangal**, Poorva Joshi, and Rahul Modak. "LSTM vs. GRU vs. Bidirectional RNN for script generation." arXiv preprint arXiv:1908.04332 (2019).