Aman Singhal

551-344-6364 | as14661@nyu.edu| LinkedIn |GitHub | amansinghal927.github.io/

EDUCATION

New York University   New York, NY
Masters of Science in Computer Science   *August, 2022 - May, 2024*
Delhi Technological University   Delhi, India
Bachelors of Technology in Electronics and Communications Engineering   *August, 2014 - May, 2018*

EXPERIENCE

UnitedHealth Group   Delhi, India
Data Scientist, Machine Learning - Awards   *July, 2021 - July, 2022*
- Collect & interpret large datasets via supervised & unsupervised deep learning, delivered artificial intelligence products
- Strengthen **AutoML** stack; added RBF SVM & reduced hyper-parameter tuning time for neural networks by 30%.
- Fine-tune Computer Vision & Natural Language Processing for **information extraction** & named entity recognition.
- Lead agile feature extraction & gradient boosting ensemble; identified class **imbalanced** PHI security restrictions.
- Won **hackathon**; developed hypothesis testing, Transformer language generation & LSTM text classification, patented.
- Drive autocomplete recommendations using k-means **clustering**, PCA reduction & Tableau business intelligence.

CVIT Lab, IIIT   Hyderabad, India
Research Assistant (Prof C V Jawahar) - Papers Code   *August, 2020 - March, 2021*
- **Published** Exploring Pairwise NMT for Indian Languages at International Conference on NLP, ACL Anthology.
- Pioneered **domain adaptation transfer learning** technique. Leaderboard neural encoder-decoder WAT **workshop**.

TransOrg Analytics   Gurugram, India
Data Scientist - Papers Code   *June, 2018 - August, 2020*
- Manage cross-functional team, gathered requirement; communicated business insights & vizualizations to stakeholders.
- Revamp market mix modeling & automated data cleaning + aggregation at **PepsiCo**; reduced 15% regression RMSE.
- Design personalized recommendation system using collaborative and content-based filtering for **Dominos**; 90% F1.
- Maintain web scraping & Argo data ingestion workflows, deployed product lifecycle management tool on IBM cloud.

RESEARCH PROJECTS

Semantic Search Engine, Contract Intelligence *Code*
- Developed YOLO object detection, segmentation and Elasticsearch + **BERT** pipeline; ranked 0.1B clauses in 90s.
- Deployed Python **RESTful APIs** with Docker & Kubernetes; optimized NoSQL **database design** for real-time search

Audio Analytics, Content Moderation (Prof Aaditeshwar Seth, IIT Delhi) *Code*
- Automated XGboost, CNN & ResNet speech identification; developed **AWS** pipelines for non-profit social platform.
- Catalyzed time-cost saving with CatBoost, synthetic data augmentation & **feature engineering**, scaled 1M household.

Multimodal Event Classification and Anomaly Detection
- Analyzed energy consumption patterns and sensor readings in smart lighting systems using audio & video time series.
- Applied logistic regression, A/B testing & experimental design; leveraged Recurrent Networks (RNN) & autoencoders.

Insurance Risk, Predictive Modeling (Montana State Fund)
- Setup Hadoop ETL &, Spark & Hive wrangling; led end-to-end exploratory data analysis & Random Forest prediction
- Collaborated with operations team and led meetings for interpretable forecasting dashboards, reduced claims 7% YOY.

Demand Forecasting, Time-Series Forecasting (Nissan Motors)
- Boosted imputation with linear regression, decision trees, KNN, neural networks, k-means & gaussian mixture models
- Implemented CI/CD, data preprocessing, outlier handling, class imbalance, dimensionality reduction, & cross-validation

SKILLS

Machine Learning: Deep learning (Keras, PyTorch, TensorFlow), HuggingFace, spaCy, Scikit-Learn, SKLearn, OpenCV, Numpy, Pandas, Data Visualization (Tableau, Matplotlib), Time-series, Forecasting models, TF Serving, Power BI

Languages & Software: Python scripting, SQL, R, Kubernetes, Cloud platforms (AWS, Azure), Docker, Git, Scala, C++, JavaScript, JQuery, CI/CD, Jupyter, APACHE, Large-scale data processing, PySpark, Hive, MySQL, LaTeX

Courses: Deep Learning, Machine Learning, Bayesian Optimization, Causal Inference, Probability Statistics, Data Structures & Algorithms, Object Oriented Programming, Multivariate Calculus, Linear Algebra, Software Engineering