ASHESH

Telephone:+886 928543959 Address:1st Floor, No. 162, Section 2, Research Institute Road, Nangang District, Taipei, Taiwan	Email: ashesh276@gmail.com Website: https://ashesh-0.github.io/ LinkedIn: https://linkedin.com/in/ashesh0 Kaggle: https://www.kaggle.com/silence2 Github: https://github.com/ashesh-0

Publication

360-Degree Gaze Estimation in the Wild Using Multiple Zoom Scales [2020] https://arxiv.org/abs/2009.06924

Accurate and Clear Precipitation Nowcasting with Consecutive Attention and Rain-map Discrimination [2021] https://arxiv.org/abs/2102.08175

Experience

Mar 2020 - Present
Research Assistant (Computer Vision)
NTU, Taipei, Taiwan

- **3D Gaze estimation** in unconstrained environments using both image and video frames as input.
  - Full 360° variation in yaw handled using sine-cosine based target space transformation. Improved prediction on frontal gazes using a weighted predictive scheme.
  - Robustness with respect to varying head sizes in images and extraction of features present at multiple magnification levels were handled jointly using multi scale feature aggregation.
  - Achieved **state of the art performance** on two datasets: Gaze360 and RT-GENE (https://arxiv.org/abs/2009.06924).
- (Ongoing) Extreme precipitation prediction for Taiwan region using Radar data. Image-to-image translation network with GRU being used currently.

Feb 2019 - Oct 2019
Data Science
Self Employed

- Participated in 4 kaggle competitions (1-1.5 month each). Was in the **top 2-3 percent** in the last 2. Details in Projects section.
- Did **5** coursera certifiable courses involving Deep learning. Details below.

Dec 2015 – Dec 2018
Data Scientist
Qplum Software Labs + Two Roads Technological Solutions, India

- **ML model** for portfolio: Development of autoencoder based market neutral strategy. Generated synthetic data to aid in training. It managed **5% of the portfolio** (Python).
- **ML model** for execution: Development and analysis of multiple intraday execution algorithms and meta algorithms. Used regularized LR

and traditional trading techniques like mean reversion, momentum. **Daily, $50K** was traded using my algorithms **saving 1-2 bps** (Python,C++).
- **ML data pipeline**: Extraction and distributed processing of data from raw tick data files and web apis. Used airflow and celery for distributed processing (Python).
- Non data science projects involved
    - Conversion of sequential simulation engine to vectorized simulation engine. Achieved **5x speedup** (Python).
    - Creation of Execution pipeline, Order routing server and Reconciliation pipeline for multiple brokers (Python,C++).

May 2015 – Oct 2015
Software Developer
Readersdoor Pvt. Ltd, Delhi, India
- Recommendation module for rooms and books. Scraping news content.

Education

B.Tech & M.Tech in Computer Science
Indian Institute of Technology Delhi
Delhi, India
July 2015
- CGPA: 8
- ML related courses: Artificial Intelligence, Machine Learning, Special Topics in AI: Probabilistic Graphical Models, Computer Vision, Digital Image Analysis, Graph Theory
- Bio related courses: Molecular cell biology, Modern biology for engineers, Biometry, Systems Biology, High Dimensional Biology, Intr. to Prac. Modern Biology

Projects

(M.Tech Project) Subcellular Regulatory Network Learning using MLN
Jul 2014-May 2015
- A model which jointly learns the biclusters and links (activating and inhibiting) in the gene regulatory network using Markov Logic Networks on Halobacterium dataset of Inferelator. Used canopy clustering results as initial state.
- With synthetic data, was able to show the limitations of our approach in terms of available data size and complexity of the network.

(Kaggle Competition) Prediction of magnetic interactions between atoms in a molecule.
Jul 2019-Aug 2019 Github link.
- Ensemble of MPNN(message passing neural networks) and GBDT. Extensive feature engineering for GBDT was done. Reached in **top 3% solutions**.

(Kaggle Competition) Predicting next month sales of products in shops.
May 2019-Jun 2019 Github link.

- Primarily feature engineering was done. Used PCA on top of TF-IDF on item names and shop names to get important features. Mean encodings, lagged features, city features and several other features were created.
- Nearest neighbors was also used to create features. GBDT was used as a model. At submission time, the solution was in the **top 2%.**

On Uniqueness of Amino Acid (AA) composition In Proteins
Jan 2014-May 2014

- Hypothesis was that average AA composition completely determines the average protein length of that class and that AA composition is unique.
- Formulated a function for predicting protein length from AA composition. Eventually ***proved*** the hypothesis otherwise.

Projects done in online courses

Guided projects done as part of coursera courses.

- **Computer Vision**: Facial recognition, Face detection, Face generation using GANs,Image captioning, Car detection with YOLO algorithm, Art generation with style transfer.
- **Bayesian Methods**: Variational Autoencoder on MNIST dataset.
- **Reinforcement learning**: Approximate Q-learning on Cartpole, DQN on Atari, Advantage actor critic on atari.

Certifiable Online courses (Coursera)

- **Done in 2019**: deep learning in computer vision, practical reinforcement learning, Bayesian methods for machine learning, how to win a data science competition, introduction to deep learning.
- **Done in 2017-2018**: neural networks and deep learning, improving deep neural networks: hyperparameter tuning, regularization and optimization, structuring machine learning projects, convolutional neural networks, sequence models.

Skills

- Areas: Deep Learning, Machine Learning, Computer Vision, Reinforcement Learning, numpy, pandas, keras.
- Languages: Python, C++, Bash, C, Matlab, ruby
- Queues: Rabbitmq, Kafka.
- Task management: Celery, Luigi, Airflow, Sidekiq, Resque.
- Database: Mysql, Neo4j, Alembic.
- Others: AWS, Elasticsearch