Akhilesh Sanjay Somani

Website: https://akhilesh-somani.github.io
linkedin.com/in/akhilesh-somani | github.com/akhilesh-somani

Education

University of Illinois at Urbana-Champaign (UIUC) May 2021
Master of Science in Mechanical Engineering **GPA: 3.91/4.0**

Indian Institute of Technology Bombay (IIT Bombay) Aug 2019
Bachelor of Technology in Mechanical Engineering; Minor in **Computer Science & Engineering** **GPA: 9.16/10**

Relevant Skills

Python, Jupyter, Data Science and Analytics, Statistics, Machine Learning, SQL, MS Excel, MATLAB, R, Kubernetes

Technical Experience

- **Data Science Intern**, *Corteva Agriscience* **May 2020 - Aug 2020**
  - Worked on predicting transgene expressions in agricultural biotech traits by implementing, tuning & comparing performances of various statistical tools **(regression)** & machine learning models **(neural networks, decision trees, ensemble methods, KNN)**
  - Built an interactive **R-Shiny dashboard** to assist biologists before conducting experiments, to save time & money
  - Used **Kubernetes** to run deep learning models on the GPU cluster for faster execution
- **Data Science Projects**, *University of Illinois at Urbana-Champaign* **Jan 2020 - May 2020**

  Autonomous Vehicle (AV) Safety Data Analysis:
  - Performed **Hypothesis testings and Kolmogorov-Smirnov tests** to compare AVs' performance results with human drivers, predicting an accident probability of 213 times higher, implying AVs not yet ready for large-scale deployments
  - Coded **Naive Bayes Model** from scratch to predict, with over 80% accuracy, causes of failure under different conditions

  Unsupervised Stool Sample Analysis in Hepatic Encephalopathy:
  - Studied stool sample data for liver cirrhosis patients to analyze & predict the microbes responsible for brain damage
  - Refined and cleaned the raw data by coding and implementing **Bayesian Networks** from scratch
  - Implemented Kolmogorov-Smirnov tests to perform statistical analysis on microbe abundance data
  - Performed dimensionality reduction using **PCA** & clustering using **KMeans, GMM, Hierarchical clustering**, etc. to successfully identify & study taxonomical relations between 20 microbes (from 150) with altered abundance levels

  Data Analytics in High-Performance Computing Security:
  - Parsed the raw data from network packets into analysis-friendly format using Pyshark
  - Identified attacker & legitimate user and DNS servers to study the origins and progression of the multi-stage attacks
  - Implemented **Hidden Markov Models** & **Factor Graphs** to predict the likely state of attacks

  Quant Analysis of Stock Market:
  - Performed feature engineering, PCA, LDA, and clustering for automated sector-identification of S&P 500 Index stocks
  - Attempted to predict future stock prices using Keras LSTMs
- **Computational Modeler**, *University of Illinois at Urbana-Champaign* **Jan 2020 - May 2020**
  - Modeled the underlying physics, and simulated & visualized the optimized motion of a slithering snake in Python
  - Coded from scratch & implemented **Genetic Algorithm** and **Co-variance Matrix Adaptation Evolution Strategy** to maximize forward velocity of snake by optimizing activation torque generated by snake's muscles
- **Summer Intern**, *University of California at Berkeley* **May 2017 - Jul 2017**

  Data Analysis of Nanoscale Thermal Protrusions on Hard-Disk Drives:
  - Used MATLAB to clean, process, and analyze data (captured from experiments on nanoscale protrusions (~10 nm) of control resistors on hard disk drives) to prove a linear relation between protrusions and power input of resistor

Key Projects

Finite Element Analysis (FEA) of Bimetallic Strip, *UIUC* Nov 2019 - Dec 2019
- Coded a MATLAB Finite Element Solver and compared results with ABAQUS, achieving errors less than 3 percent