(469)-525-5785
Plano, TX

Arjun Chintapalli

arjun.ch@gatech.edu
[U.S. Citizen]
github.com/arjunchint

Education

Georgia Institute of Technology — ***Online M.S., Computer Science*** — **December 2018**
Georgia Institute of Technology — ***M.S., Computational Engineering*** — **December 2017**
University of Texas at Austin — ***B.S., Petroleum Engineering Honors*** — **May 2016**

Certifications: AWS Big Data, AWS Solutions Architect, AWS Developer, AWS SysOps, Engineer-in-Training
Courses: Machine Learning, ML for Trading, Reinforcement Learning, Algorithms, Numerical Linear Algebra, Parallel HPC, Simulation, CFD, Machine Vision, Data Analytics/Visualization, Info Security, Big Data w/ Spark

Experience

- **Data Engineer** - *Capital One* — **January 2018 - Present**
  - Setup AWS infrastructure to implement cloud/Spark migration of Ab Initio ETL pipeline
  - Deployed SQL recommendation and monitoring engine to reduce Snowflake compute costs using AWS Lambda
  - Created enterprise data pipeline that matches/decrypts/renames files landing in S3 data lake and triggers ETL jobs
  - Released process to stream S3 Read/KMS Decrypt/Write +100 GB files from data lake using S3 Events/Lambda
  - Setup regex catalog of 1000+ file patterns to match/capture/rename incoming S3 files and trigger relevant jobs
  - Implemented PiT DR failover/recovery by tracking processed files using SNS, DynamoDB and Lambda
  - Released SSN/PCI scanner to scan outgoing files/data directories periodically for sensitive data
  - Automated ASG deployment, software installation and data recovery using CFT's, Lambdas, Snapshots
  - Created Lambda to EC2 snapshot backup, tag and copy to DR and EC2 restoration/mounting scripts
- **Drilling Data Intern**- *Intellicess* — **January 2017 - August 2017**
  - Developed a real-time MongoDB cloud database to provide drilling metric visualizations from rig data
  - Improved Bayesian rig state classifier and washout belief predictor using real time sensor data
  - Developed a well completions recommendation system based on lateral downhole MSE data
  - Created real-time drilling efficiency metrics and rig parameter recommendation system to fasten drilling
- **Reservoir Engineering Intern** - *Raisa Energy* — **June 2015 - August 2015**
  - Compiled well production, entity and downhole data from databases for decline curve analysis
  - Modeled impact of factors like operator, date, lateral length, location on profitability metrics
  - Determined optimal drilling areas by analysis of net present value, decline rate and recovery
- **Drilling Data Research Assistant**- *University of Texas at Austin* — **May 2014 - May 2015**
  - Created Bayesian belief network to process drilling sensor errors and generate alerts
  - Integrated calculated features such as Mechanical Specific Energy (MSE) to real time drilling data
  - Automated generating daily rig reports evaluating daily performance and generating recommendations

Projects

- **Spark GraphX**: Scala Spark graph analysis to find similar patients using PageRank, power iteration, and phenotyping
- **Spark MLlib**: Predicted drug recall using NLP sentiment features with CNN/RNN model and PySpark pipeline
- **Deep Q-Learning**: Recreated Deep Mind DQN net to win Atari games, analyzed performance w/ varying architecture
- **ML Trader**: Developed market trading algo utilizing Q-Learning and market features such as volatility, momentum
- **CUDA CFD**: Created GPU parrallized lattice fluid simulation using CUDA, optimized using NVIDIA profiler
- **HPC**: Implemented multi-node distributed bucket sort using OpenMP and MPI library on HPC cluster using Slurm
- **NLP**: Used Tweepy API, Beautiful Soup and NLTK to cluster relevant tweets and predict 2016 election sentiment
- **Cybersecurity**: Implemented programs that accomplish CSRF, XSS, and SQL injection attacks
- **Computer Vision**: Conducted image analysis using neural nets, PCA and Kmeans to identify diseased food products
- **Embedded Smart Home**: Created iOS app to control lights, HVAC and touch-screen display using Arduino
- **Time Series Analytics**: Conducted PCA, Fourier and ARIMA analysis on time-series voltage data to predict outages
- **Reservoir Simulation**: Reservoir simulators, fluid flow/properties, field project economics, value of info of sensor data

Accomplishments

Co-author of SPE Paper, "Self-Learning Probabilistic Detection and Alerting of Drillstring Washout" — May 2018
Co-author of SPE Paper, "A Novel Probabilistic Drilling Optimization Index" — May 2017
Recipient, UT Petroleum Engineering Department Scholarship — 2013 - 2016
Recipient, National AP Scholar [Maximum Scores on 15 AP Exams] — June 2012