Cynthia Hanson Singh

chansonsingh@yahoo.com | 281-740-1637 | www.linkedin.com/in/cynthia-hanson-singh-27783310b | Houston, TX 77070

Profile:

- Data scientist/analyst and engineer along with strong mathematical ability seeking a data scientist/mathematical position as statistical scientist and machine learning engineer and developer of complex algorithms and software developer with Python and R, also MATLab to handle data with solutions to complex analysis.
- Equipped in Python and R programming and writing algorithms to analyze, clean and wrangle data, as well as handling data using PostgreSQL, SQL, NoSQL such as Neo4j and MongoDB.
- Experience building ML/AI models with Python and R (statistical use) for Deep Learning, NLP, Computer Vision
- Experience with use of algorithms for ML/Deep Learning/AI projects
- Experience with PySpark, Spark, writing SQL queries and analysis of big data.
- Experienced implementing Kafka for steaming data and moving data from system to system
- Skilled in C++ programming.
- Experience with cloud-based applications such as AWS and Azure and Data Bricks, Data Lakehouse with Azure.
- Programming experience with numerical Python and handling data with Python. Experience with Seaborn and Matplotlib for graphical results.
- R programming experience with statistical analysis and graphical packages.
- Implemented PyTorch and TensorFlow for Deep learning and data visualization
- Paralleled with being a highly trained electrical engineer.
- Implement and use ANOVA, GLM, ANCOVA models and methods to analyze data
- Developed and use LME (mixed effects) models, with correlated and autocorrelated data
- Progressive thinking and problem-solving is used, detail-oriented ability implementing engineering knowledge and background, and strong interpersonal skills along with a special ability communicate verbally and in writing effectively with employees is necessary.
- Create and modify data reports and dashboards for sharing data securely across applications with cross-functional teams using AWS and Power BI
- Motivated and driven to lead multi-disciplined or cross-functional teams involved in scientific or engineering development programs/projects.
-

Computer Skills:

- Python, C, C++, Pyspark, MATLAB, R, LabView, SQL, SAS, AWS Cloud Platfor , Mathcad
- Mathematica, Spice, PSpice, SRIM/TRIM, Sylvaco, Microsoft Office suite are some tools of proficiency.

Professional Experience:

Steampunk   **Aug 2023 to Present**

Data Engineer/Data Scientist

Client: USDA

Responsible for data migration of big data to AWS cloud platform. Several mainframes and servers used by USDA, which are becoming obsolete, are becoming replaced by SQL, PySpark, RDS, Redshift, Tableau, and Athena.

- Data analyst responsibilities include analysis of all data from all servers and mainframes to be migrated to AWS platform. PySpark scripts, Glue scripts, Postgres conversion scripts written to import data.
- Drivers and client servers installed and connections made from original servers to AWS cloud platform.
- Validation functions created, optimization and testing scripts created and run prior to migration.
- DBeaver implemented to import databases and schema. Ran queries to confirm successful import.
- State machines, Lambda functions and Glue scripts built for migration of data in AWS.
- Collaborated with stakeholders to ensure ETL/ELT migration plan is desired and in agreement.
- Incorporated raw data analysis with CI/CD implementation and API analysis conducted and endpoints within AWS built and constructed for big data migration.
- Transformed and analyzed SQL scripts to be migrated from one source to AWS.

Eva McKenna Consulting, LLC   **Jan 2019 to May 2019 2023**

Data Scientist/Analytics/ML/NoSQL

Client: Outdoor Recreational Product Producer

Designed and developed algorithms and regression models for cost analysis for the development of new and advanced equipment and products. Performed regression and ML train/test set analysis for predictive forecasting of sales using Python and R.

- Collaborated with business partners and business analyst regarding desire of data analytics and data output.
- Prepared and used churn analysis to increase sales, predict areas of improvement in business analysis
- Prepared raw data using data mining and wrangling techniques by collecting, cleaning, and analyzing data.
- Incorporated machine learning algorithms, statistical analysis, regression, and forecasting algorithms
- Data preprocessing techniques specific to image/video data, including augmentation and normalization
- Implemented sampling techniques such as random, systematic, clustering, classification sampling.
- Incorporated machine learning algorithms and advanced statistical analysis such as random forests, regression and classification models using NumPy, pandas, SciPy, sklearn, LASSO from Python

- Building and deploying image/video analytics models at scale using cloud platforms (e.g. Azure, AWS)
- Work in cross-functional teams to provide clients, detailed visual data reports with EXCEL and Power BI
- Built dashboards and data reports using Power BI
- Analyzed data consisting of univariate, bivariate, and multivariate analysis
- Implemented ML analysis with AWS cloud application
- Implemented data visualization with Seaborn and matplotlib and R graphical packages
- Analyzed and wrangled data using MongoDB in NoSQL
- Utilized Neo4j to incorporate data summaries through topic modeling algorithms that uncovered related tags in data
- Implemented SAS to structure data according to dates, ID records, comparative analysis for cross validation.
- Implemented CI/CD protocols
- Used Kafka to build real-time streaming data pipelines and real-time streaming applications

Client: Cement Products   May 2019 to Jan 2021
Data Scientist/ML Analysis

Developed time series analysis of employee time reporting, implemented a product/cost analysis of data, set up data and app containerization, designed hands-on ETL processing.
- Use Python, SQL, PostgreSQL for data analysis, implement Python libraries NumPy, Pandas, Seaborn, Matplotlib
- Achieved SAS to Python translation implementation projects.
- Incorporated GLM and ANOVA and ANCOVA models for statistical analysis
- Develop and deploy deep learning models for image/video analytics applications
- Educated and demonstrated to cross functional teams implementation of data visualization using Power BI
- Incorporated using Data Bricks Lakehouse Platform for analytics and AI analysis
- Deployment for use with API, PySpark for Apache Spark, and use of ML lib.
- Experience in TensorFlow and PyTorch for ML analysis using Python
- Incorporated Amazon Web Service (AWS) cloud platform for developing and testing different predictive models.
- Implemented Apache Spark projects to develop open-source software for reliable, scalable distributed computing.
- SQL and PostgreSQL utilization to organize, join and sort data according to business structure and needs
- Utilized AWS for ETL purposes for storing and sharing data among cross-functional teams.
- Deployed ETL data loading both full and incremental loading
- Combined R and Python for data analysis to develop trends and hypothesis testing.
- Led of team of 3 to demonstrate data extraction, data analysis, machine learning execution in Python, R, SQL

- Developed project using Python Flask for web application
- Incorporated the use of Docker and GitHub within the agile team
- Used Kubernetes for automating deployment, scaling, ad management of containers

Client: Cosmetics Company   Jan 2021 to July 2023

Data Scientist /NoSQ/SQL

Incorporated Python packages and R statistical analysis to analyze business costs, sales, and marketability for clients' needs and forecasting spectrum. Implemented NoSQL Neo4j to create relational graphs to improve sales and marketing. Oversaw data visualization goals to meet business collaborators' needs.

- Used ML packages implemented in Python, NoSQL application to decipher relationships between customer and product, developed ETL structures for data flow to cross-functional teams
- Collaborated with cross-functional teams to explore data from SQL queries and organization of data with SAS.
- Implement Python, API, REST API, with web services AWS and Azure for customer analytics.
- Use of modeling techniques and machine learning to improve analysis prior to presentation to client.
- Develop and deploy deep learning models for image/video analytics applications.
- Incorporated AWS Cloud Platform to store, secure, move, and extract data.
- Analysis completed with Python scikit library, Python algorithms and functions.
- Use of R programming to organize, clean and analyze data.
- Open-source libraries for computer vision tasks, such as OpenCV:
- Performed cluster analysis to identify most marketable sales and profit margins.
- Utilized ETL methods to clean and transform data to consolidate raw data into staging area.
- Joined and split data for munging purposes using relational data base management systems: SQL and SAS.
- Implement use of NoSQL database using MongoDB and Neo4j applications.
- Utilized Neo4j to analyze relation between specific products and customers' influential markers.
- Implemented ID records, relational interests within database records.
- Designed and implemented custom, proprietary visuals using Power BI

Lonestar Montgomery, TX   Aug 2013 to Dec 2018

Adjunct Professor of mathematics and engineering

- Taught the introduction to engineering classes, and both statics and dynamics in mechanical engineering classes.
- Taught college algebra classes among others.

TcSAM

Jan 2007 to Dec 2010

Electrical Engineer/Research Engineer

- Developed the control of MBE vacuum system using C++ and NI LabView.
- Developed, devised simulation and modeling testing techniques for research groups with MATLAB and Python as per request.
- Oversaw laboratory practices and instilled clean methods.

University of Houston   May 1997 to Dec 2007

Research Assistant

- While working on research for my PhD, excelled at performing analyses and returning feedback in a timely manner to professors or other graduate students within the research group programming with LabView from National Instruments and C++ to control electrical equipment that involved gases, vacuum pumps, depositors, sensors, motors, high temperatures; using many different tools but not limited to SEM (scanning electron microscope), MBE (molecular beam epitaxy), Normarski microscope, depositing metals using the thermal and e-beam evaporators, performing photolithography on samples for patterning, writing LabView programs which controlled the use of the MBE which includes eurotherms, stepper and servo motors, Riber shutter controllers, and EPI valve positioners, and CLR measuring equipment; cleaving samples, RTA (rapid thermal annealer), ran modeling and simulations on a test structure before experimental procedures were done on it using Sylvaco software and SRIM/TRIM.
- Proficient using MATLAB and LabView for material science studies on various substrates.

University of Houston   Aug 2000 to May 2001

Teaching Assistant

- As a T.A., graded undergraduate electrical engineering courses and taught in an electrical engineering laboratory used by all undergraduate students.

University of Houston   Aug 1996 to Aug 1999

Undergraduate Assistant in Space Physics

- Detailed in the development of circuit layouts for supervising professors using ORCAD, making the circuit boards and soldering the components onto circuit boards.
- Instrumental in the orders for all the needed electrical components for future circuit board designs through many vendors.

Education:

- Texas A & M University 2019 to Present College Station, TX (will complete Dec 2023)
- Currently working on and completed 32 hours of **Master's mathematic/statistics** classes including:
- Statistical Courses (using Python, R, SAS, SQL, Docker, C++, PySpark, Spark as examples)
- Big Data Analysis and tools for analysis to ensure data integrity.
- Regression Analysis using R.
- Modeling Deep Neural Networking Systems such as CNN, RNN, and LSTM
- Dynamical Systems
- Methods in Applied Partial Differential Equations
- Linear Algebra for Applied Mathematics
- Mathematical Modeling

AWS Certification: Presently working on:

- Data Science and Business Analytics master's Program
- AI and Machine Learning

University of Houston Jan 2002 to May 2007 Houston, TX

- ABD Doctoral program in Electrical Engineering within the optoelectronic engineering research group at TcSAM (Texas Superconductivity and Advanced Materials), which is a company that works in material sciences.
- Completed all classes and research.
- Cumulative GPA 3.9

University of Houston Aug 1999 to Dec 2001 Houston, TX

- M.S., Electrical Engineering within the Microelectronic research group
- Cumulative GPA 3.85

University of Houston 1993 to 1997 Houston, TX

- B.S., Physics

Relevant Courses in MS and PhD:

- Databases and computational tools used in Big Data (using Python, R, SAS, SQL, PySpark and Spark)
- Statistical computations (using Python, R, MATLAB)
- Advanced Digital Design: experience with Verilog programming
- Introduction to Stochastic Processes
- Digital Logic Design
- Advanced Process Integration for VLSI (Very Large-Scale Integration)

Achievements:

- Received a one-year fellowship through Texas Space Grant Consortium 2005-2006 for the work I was doing designing a VLWIR (very long wavelength infrared) photodetector.
- Served on the Board of Directors for the Cy-Fair Swim Club.