AVINASH YEKKALA

[avinash.27291@gmail.com](mailto:avinash.27291@gmail.com)

https://[www.linkedin.com/in/avinashyekkal](http://www.linkedin.com/in/avinashyekkala)a

857-308-9600

Data Scientist and Engineer with 9+ years of experience in developing machine learning and deep learning models, predictive analytics, forecasting, statistical modeling, data processing and visualization to solve business challenges. Well-versed in working with Python and R programming languages. Able to carry out various scope of research projects from inception to production, including the interpretation of complex data. Experienced in leading a team and working with cross- functional teams. Built a chatbot from scratch using NLP techniques and maximized ROI. Proven experience in below domains

* Insurance
* Banking/Finance
* Airlines

**TECHNICAL SKILLS**

**Programming Languages**: Python, SQL, R

**AI Libraries:** TensorFlow, Keras, NLP **BI Tools:** Tableau, PowerBI, Qliksense **Cloud:** Azure, AWS, Oracle cloud

**ETL Tools:** SSIS, Alteryx, Informatica.

**Big Data:** Hadoop, Pyspark

**Web Analytics:** Google Analytics

* Healthcare
* Retail

**PROFESSIONAL EXPERIENCE**

**Google, Mountain view, CA**

**Lead ML Engineer– Generative AI March 2023 - Present**

* Summarized conclusions of Business Intelligence Analysis and presented findings at executive meetings.
* Generated daily, monthly and annual reports to address short or long term business questions.
* Owned Youtube community posts project and developed sql queries and tables to get the required data and thereby deploying the sql to production.
* Have experience working on Google cloud platform (GCP). The services that I used for my Data science work are Google Colaboratory, DataLab, Vertex AI for data science and generative ai purposes.
* Owned world wide blog migration of Google Analytics to latest version 4. Migrated all the 10 blogs which our team maintains to Google analytics 4.
* Set up tag manager related to Google analytics and developed reports on website clicks, outbound and event name etc to populate in BI reports
* Created documentation related to GA4 and explained the process of migration and setup of tag manager.
* Partnered with account executives to execute the AB testing(control and treatment) groups for Twitter/Youtube posts.
* Planned agile scrum sprints and managed a team of 4 related to GA4 and trained them on the new process of web analytics migration.
* Developed Machine learning to eliminate the need to use bug management tool and predicted the classes of metadata.
* Automated the requirement of manual intervention of updating google docs related to google publications by using machine learning LLM Generative AI models.

**Alight Solutions, Chicago**

**Lead ML Engineer (NLP) – Generative AI March 2022 – Dec 2022**

* Owned and Analyzed business use cases and extracted relevant data from Apache Hue Prod by developing SQL Queries.
* Responsible for understanding data and developing sql queries and ML algorithms and deploy into production
* Developed Generative AI NLP LLM code to find the relevant keywords for different health care units such as Child care, Elder care etc and extracted more relevant data using those keywords from our big data for data science purpose
* Used various machine learning algorithms such as regression, classification and NLP etc.., to predict the chatbot data and health care insurance data for various business use cases and optimized it with optimization algorithms.
* Deployed ML algorithms using MLOPS AWS environment such as step functions, lambda, S3 and pickle files.

**Microsoft, Seattle, WA Apr 2020 to Dec 2021**

**Data Scientist (Natural Language Processing) – Conversational AI**

* Analyzed Textual and structured data and solved intricacies of existing data warehouse which ultimately led to solve ML problems (As I use this data for developing ML chatbot).
* Brainstormed with my manager and created SIX different ideas which are non-existent in my project and prepared a roadmap for my ideas.
* Lead onshore and offshore team and worked in different time zones and communicated my ideas to them and prepared Python-Machine Learning, ETL and SQL code POC’s for my ideas.
* Owned and created scenarios for conversational chatbot using NLP and developed Natural Language processing ML algorithms to train chatbot (for example Topic modelling, Deep Learning) to generate answers automatically to customer questions in real time in databricks environment.
* Developed ETL pipelines and Data science pipelines using Azure databricks and deployed it into production.
* Created visualizations using databricks which showed the impact of my ideas and how much Microsoft saved in dollars with the help of chatbot (Microsoft saved 356 thousand US dollars in two quarters) by reducing human dependency.

**Axtria, Boston, MA 2018-2020**

**Data Scientist**

* Involved in all phases of data acquisition, data collection, data cleaning, model development, model validation, and visualization to deliver data science solutions.
* Performed data analysis and developed ETL pipelines for multiple marketing departments which helped my team in decision making.
* Developed advanced Data science models using python to predict the company’s health care business.
* Built key business metrics, Visualizations, dashboards, reports with tableau.
* Performed regression analysis on data sets using linear models to extract insights and forecast trends.
* Worked with all levels of stakeholders to develop reports, used for quality reviews and other desired outcomes

**SapientRazorfish, Atlanta, GA 2018 -2018**

**Data Engineer – Strategies & Advertisements**

* Designed, modeled, validated and tested statistical algorithms against various data sets including behavioral data and deployed predictive models.
* Created SSIS ETL jobs and used machine learning algorithms for price predictions.
* Integrated the data and migrated it to Amazon web services and maintained data in cloud.
* Performed Data Collection, Data Preparation, Feature Engineering, Hypothesis Testing, Dimensionality Reduction and Data Mining to help Data Scientists developing mathematical and statistical models.
* Collect comprehensive and accurate data from the point-of-sale systems. This includes transactional data such as sales, items purchased, timestamps, customer information, and any other relevant variables. Ensure the data is clean and properly formatted.
* Choose appropriate modeling techniques based on the nature of the problem and the available data. Common techniques used in POS modeling include regression analysis, time series forecasting, clustering, classification, and recommendation systems.
* Constructed Natural language processing (NLP) code to analyze unstructured data and performed sentiment analysis of the customer by using customer call data.
* Developed deep learning models POC’s using python to predict the outcome for various datasets.

**Kony Labs, Hyd, India 2014 - 2015**

**ETL Test Engineer**

* Designed and developed ETL test cases using business rules document and executed ETL test cases and updated the results to defect management tool.
* Performed data analytical testing for Business Intelligence systems and validated data transformations and performed end-to-end data validation for ETL and BI systems.
* Tested Informatica ETL data jobs which were failing and debugged ETL pipelines using error logs, which resulted in uninterrupted data flow process to target systems.
* Worked on SQL Queries which are taking higher processing time and performed performance tuning which resulted in reducing the processing time of data.
* Created automated test scripts using python and fetched data automatically to validate data.

**C-Edge Technologies, Mumbai, India 2012-2014**

**Assistant System Analyst**

* Performed Data Analysis on customer financial transaction records using python libraries.
* Researched and explored suitable machine learning techniques based on data sets to predict potential business.
* Worked on customer churn model to predict the credit card trading business and identified threats related to customer attrition, potential defaulters which causes business loss by executing machine learning algorithms like regressions and classifications there by improving profits by 2%.
* Created story-board on business discoveries using Business Intelligence tools and communicated to end clients which resulted in actionable outcomes
* Performed Data Collection, Data Preparation, Feature Engineering, Hypothesis Testing, Data Reduction and Data Mining to help Data Scientists developing mathematical and statistical models.

|  |  |
| --- | --- |
| **EDUCATION** |  |
| Masters in Information Systems, Northeastern University, Boston, MA | 2016-2017 |
| Bachelors in Electronics and Communication Engineering, JNTUH | 2008-2012 |