

# Arthkumar Patel



Mission Viejo, California



(949) 527-8145



arthp7044@gmail.com



<https://arthpatel98.github.io/>



<https://www.linkedin.com/in/arthp98/>

## EDUCATION

Master of Science, Computer Science

**San Diego State University, San Diego, CA**

( Jan 2021 – Dec 2022 ) ( GPA: 3.92/4.0 )

Bachelor of Engineering, Computer Engineering

**LDRP - Institute of Technology & Research,**

**Gandhinagar, INDIA**

( Aug 2016 – June 2020 ) ( GPA: 8.58/10.0 )

## RELEVANT SKILLS

- **Programming Languages** - Python, Java, C, C++, HTML5, CSS3, JavaScript, Shell Scripting (Linux), PHP
- **Machine Learning Frameworks** - TensorFlow, Scikit-learn, PyTorch
- **Data Analysis & Visualization** - Numpy, Pandas, Seaborn, Matplotlib, Tableau, Microsoft PowerBI
- **Web Frameworks** – Angular.JS, React.JS, Flask, jQuery, Node.JS, REST
- **Databases** - SQL, MySQL, NoSQL, MongoDB
- **Platforms** - Git, AWS EC2, JIRA, Docker, Apache Spark, BitBucket, PySpark, Android Studio
- **Other** - Agile Frameworks, System Design

## CERTIFICATIONS

[Google Data Analytics Specialization](#)

**Learnings:** Data Exploration, Data Cleaning, Data Analysis, Google Data Analytics Capstone Project

[IBM Data Science Specialization](#)

**Learnings:** Data Science Tools, Python for Data Science & Artificial Intelligence, Databases & SQL for Data Science, Data Visualization with Python

[Stanford University ML Certification](#)

**Learnings:** Supervised Learning, Unsupervised Learning, Regression, Back Propagation, KNN, Neural Networks, Recommender Systems, SVM

[Udemy Tableau Certification](#)

**Learnings:** Charts, Join, Blend, Calculated Fields, Dual Axis, Data Extracts, Dashboard, Level of Detail, Time-series, Parameters, Storylines

## LANGUAGES

- English - Proficient
- Hindi - Native
- Gujarati - Native

## CAREER OBJECTIVE

Curious and creative mind, willing to strive for solving problems relevant to the realms of Data Science and Software Engineering. 1+ years of experience as a Data Science Professional and Software Engineer focused in system improvements and Data Engineering. Passionate for developing innovative programs that expedite the efficiency and effectiveness of organizational success and find insights from data to achieve business goals.

## PROFESSIONAL EXPERIENCE

Data Science Intern	RipeMetrics, San Diego, CA ( Jan 2022 – Present )
<ul style="list-style-type: none"><li>➤ Performed Data Analysis on large customer datasets to extract concrete insights using advanced Python Libraries, make data-driven decisions, build &amp; test Supervised and Unsupervised ML models.</li><li>➤ Used datasets to drive optimization and perform Data Cleaning &amp; Preprocessing using Pandas, Numpy, Statistics and created effective and interactive data visualization charts in Tableau. Structured data queries and data mining through MongoDB database.</li><li>➤ Created Product Recommendation Model using Scikit-learn, PyTorch and classified customers into discount groups using K-Means Clustering and Jupyter Notebook. Implemented Market Basket Analysis Model to create product baskets based on customer buying patterns using MySQL and Apriori Association Ruled Algorithm. Architected and built data pipelines and batch jobs using Python and HDFS.</li><li>➤ Used AWS services like S3, Redshift, and Tableau for automating capturing, processing, analyzing and visualizing data developed SQL codes to process and update backend data in Amazon Redshift.</li><li>➤ <b>Technologies Stack</b> - Python, SQL, Spark, Scikit-Learn, K-Means, SVM, MongoDB, HDFS, Jupyter, Tableau, PyTorch, Numpy, Pandas</li></ul>	
Data Analysis & Digital Media Assistant	SDSU Global Campus, San Diego, CA ( July 2021 – Jan 2022 )
<ul style="list-style-type: none"><li>➤ Performed Data analysis using Python libraries, Database Monitoring &amp; Management using SQL on the data gathered from digital media marketing campaigns of SDSU &amp; visualized that data using Tableau to find students conversion ratio.</li><li>➤ Assisted in Paid Advertising Campaigns running on Google Ads &amp; Facebook Ads to perform search term reporting &amp; website exclusions and created detail documentation and reports with key findings and recommendation strategies for improving student activities to achieve goals.</li><li>➤ Performed descriptive statistics, trend line analysis, percentage point gap and proportion to evaluate student success metric using Advanced SQL, Python and ETL Data pipelines using Python, Airflow for Data extraction and Manipulation from Salesforce.</li><li>➤ Implemented &amp; supported reporting and analytics infrastructure for internal business customers using AWS, services such as Athena, Redshift and EMR.</li><li>➤ <b>Technologies Stack</b> - Python, Google/Facebook Ads, Redshift, Tableau, MySQL, Hadoop, Athena, AWS</li></ul>	
Software Engineer Intern	WaytoWeb Pvt. Ltd, Ahmedabad, India ( June 2019 – Nov 2019 )
<ul style="list-style-type: none"><li>➤ Identified and translated customer requirements into detailed architecture and designed system, created APIs (using Flask) that can ingest and process data from various sources.</li><li>➤ Developed test scripts for automation and performed unit testing of APIs and other services. Implemented some python libraries for REST API command line interface. Also, devised an AWS Batch based architecture to parallelize analysis in the analytics tool, which will decrease python runtime scripting for data validation.</li><li>➤ Developed a backend application using Node.js and Express to handle requests from web clients. Also, developed highly scalable end to end application to process 70,000 real time requests per second (Even in the case when main frame servers are down).</li><li>➤ Ensured Unit, Integration tests, Code reviews, Continuous Integration &amp; Deployment using Jenkins, Git and Agile Scrum practices for efficiency. Also, fixed issues, completed unit, regression and system test for web and desktop application ensuring consistency throughout development and improvement phases.</li><li>➤ <b>Technologies Stack</b> - HTML5, CSS3, JavaScript, Angular, NodeJS, Git, Python, MySQL, MongoDB, Agile</li></ul>	

## PROJECTS

### NBA Games & Players Data Analysis ( July 2021 - August 2021 ) [Link](#)

- By using NBA data between 2003-2020, performed analysis and visualized some of the results using pandas, numpy, matplotlib and interactive python functions.
- Created players and teams performance Comparison models, answered set of questions about players' performance and teams' performance by my analysis. Built Predictive models like Regression and Classification to predict performance.

### 2020 US Presidential Election Donations Analysis ( May 2021 - June 2021 ) [Link](#)

- From the dataset of more than 30 million donations, I analyzed how much money each candidate received in 2020 US Election, also finding mean, median of donations for each candidate, unique contributors, monitoring small contributors and their impact on Presidential Campaigns.
- Implemented code locally using Spark Shell and on cloud using AWS EC2 which provided faster implementation of Dataframes using SparkSession from pyspark.sql

### US COVID-19 Data Analysis ( February 2021 - April 2021 ) [Link](#)

- Performed analysis on COVID-19 cases, deaths and vaccines datasets for all counties and states in the United States. Also, created an interactive dashboard for my results.
- Also found out how vaccines impacted on new cases, deaths percentage and whether any state or county reached to herd immunity using numpy, pandas, seaborn and matplotlib.