

Contact

satyambpandit@gmail.com

www.linkedin.com/in/satyam-pandit-85b44182 (LinkedIn)

Top Skills

Generative AI

AI Agents

Retrieval-Augmented Generation (RAG)

Languages

Marathi (Professional Working)

English

Hindi

Certifications

Databricks Certified Generative AI Engineer Associate

Databricks Certified Associate Developer for Apache Spark 2.4

Talend Data Integration v6 Certified Developer

Honors-Awards

Analytics Interactive Mobility Achievers Award

Satyam Pandit

Databricks Certified Associate Developer for Apache Spark
Greater Chicago Area

Summary

DATA ENGINEERING:

1. ETL: Talend (Certified Developer)
2. PySpark (Databricks Certified Developer for Spark)
3. Python

DATA ANALYTICS:

1. Big Data: Hadoop MapReduce, Apache SPARK, Hive, HDFS
2. AWS Technologies: EMR, EC2, S3, Redshift, Glue Catalog, Athena
3. Data Mining:
 - a. NLP using DialogFlow, NLTK
 - b. Classification using CHAID, CART, Bagging, Random Forest, Linear, and Logistic Regression
 - c. Clustering using K-Means, KNN, Hierarchical Clustering
4. Business Data Visualization: Tableau, D3js, QlikView, Power BI
5. Statistics: Descriptive Analytics, logistic regression, ANOVA, correlation, T-Tests, chi-square testing

SOFTWARE DEVELOPMENT:

1. Enterprise Application Development: Web Application Development using Angular2, Mobile Application Development using React-Native, NodeJS, PrimeNG, JSP, JSF, Javascript
2. Java full-stack development

Experience

Wavicle Data Solutions

Data Engineering Architect

April 2018 - Present (7 years 7 months)

Greater Chicago Area

Cloud Migration Project: Key participant in data migration project from on-premises Cloudera Hadoop system to AWS cloud, to reduce project costs and

improve performance. Developed Talend jobs, Python and PySpark scripts for data ingestion and integration.

Pricing Snapshot Project: Designed a data warehouse to load Pricing Snapshot data of a chain-restaurant per store per date from the source XML files residing in S3 bucket/ HDFS and store in Redshift database to analyze the sales of the products based on price change

Natural Language to SQL: Developed a POC using Python Flask and DialogFlow that transforms natural language questions into corresponding SQL queries for relational databases

Nutrition Calculator: Developing a mobile application as part of a POC to provide food items of nearby restaurants based on price and calorie limit per meal.

FCB Chicago

Project Intern

September 2017 - December 2017 (4 months)

Greater Chicago Area

Designed an efficient data structure by data cleaning & ETL operations for Simmons data

Implemented data models & propensity score matching to determine likelihood, patterns among different user groups and visualized key results

Performed competitive analysis for a leading food retail channel to increase sales

Skills:

R | Excel Power Query | VBA | Power BI | Tableau | D3js

HERE Technologies

Engineering Intern

June 2017 - August 2017 (3 months)

Chicago, Illinois

Designed and developed a geospatial web application in a test-driven environment, integrating Elastic Search; providing access to metadata of HERE assets

Upgraded to RESTful web services from SOAP to efficiently retrieve metadata from MongoDB

Automated and integrated the process of user feedback in the web application
Enhanced existing application by adding features based on user request in an agile environment
Delivered technology solutions as part of a Scrum team in an Agile fashion including sprint planning, complexity analysis, standup calls, and sprint retrospectives

Skills:

AngularJS | TypeScript | NodeJS | PrimeNG | Elastic Search | Web Services - REST | Gerrit | JIRA

TapGenes

Project Intern

February 2017 - April 2017 (3 months)

Chicago

TapGenes helps in tracking the family health story by building a family health history tree.

The UIC team was assigned to perform challenging tasks that were successfully completed by the team.

1. Developed risk assessment models for the following-
 - a. Diabetes Risk Assessment
 - b. Lifestyle Wellness Risk Assessment
2. Created unique, appealing and user-friendly visualizations to go with the assessment that would display the results of the assessment.
3. Suggested a scoring model for Lifestyle Wellness Risk Assessment.
4. Updated the Diabetes Risk Assessment model using swagger on the UAT website.

Accenture

Software Engineering Analyst (Information Technology)

November 2013 - June 2016 (2 years 8 months)

Mumbai Area, India

I have worked with Accenture as a Software Engineering Analyst. During my stint at Accenture, I was technically trained in Java and had worked for a year on developing Java based desktop applications. I was also involved with Data Warehouse and Business Analytics team where I worked on various Data visualization tools like Tableau, D3.js and Qlikview. I have developed multiple assets and dashboard reports based on client requirements using the above tools.

Education

University of Illinois at Chicago

Masters of Science, Management Information Systems · (2016 - 2017)

University of Mumbai

Bachelor of Engineering (B.E.), Computer Engineering · (2009 - 2013)