Pratik Vagyani

Software Development Engineer

📱 **715-430-2274**

✉ pratik.vagyani2nov@gmail.com

in linkedin.com/in/pratik-vagyani

○ github.com/pratik027

PROFESSIONAL EXPERIENCE

Symantec,India | Software Engineer II   Jan 2018 - Jun 2018

Digital Certificate Classifier: Designed & developed applications for **Spark** application to process millions of records & classify malicious certificates which **saved 200 hrs/week** of human work. Spark app reliably process **10TB** data daily. Optimized spark queries upto **30%**. Performed integration & unit tests. Used Kafka queues to publish results.

Java, Hadoop, Spark, Hive, Oozie, Data Analytics, Shell, Linux

Symantec,India | Software Engineer I   Jul 2015 - Dec 2017

Signer Automation: Implemented **Master-Slave** architecture on AWS to automate signer processing at large scale. EC2 auto-scaling & SQS used to operate cost effectively **(25%)**. Developed Python scripts to update scanner and email daily stats. Performed additional role of **scrum master** for Agile team.

Python, AWS, SQS, Kafka, Unit-test

Feed Efficacy : Developed Django, Bootstrap based website to visualize feed efficacy with **graphs & charts** as well as an **alert system** to notify anomalies in feed. Automated data sourcing and storing in uniform format from MS-SQL, Postgres DB & Hadoop to improve html rendering by **60%**. Oozie, Scoop, Hive is used for data sourcing and stat generation.

Python, Django, Oozie, Sqoop, Hive, PostgreSQL, MSsql

Symantec,India | Software Engineer Intern   Jan 2015 - Jul 2015

Built a honeypot system to allure hackers and malware spammer to source the latest threats and malwares to save customers from **zero-day attacks**. Used Lamp stack, Bootstrap to develop the website and python automation to process and analyze collected malwares and threats.

PROJECTS

ExtraShift — *Distributed publish–subscribe*

Implemented distributed pub-sub system to solve the issue of open shifts by connecting employees and campus restaurants. System is built with a network of 4 dockers. Established communication in dockers by socket programming. MySQL database is used for data backup to avoid data loss incase of node failure.

Python, Docker, Socket, flask, JavaScript, MySQL

Folder Path Clustering   — *Big Data App*

Research and built app for classifying malicious folders, based on files it contains. Compressed 6 month data with bucketing, parquet /orc file formats upto 30% for faster and efficient data analytics. Performed debugging & testing of python, java & Hadoop(spark, hive) applications. Scheduled Oozie job to run Hadoop apps.

Hadoop, Hive, Spark, Oozie, Linux

WeatherOnMyTrip

Built a website to help travelers to visualize weather on trip route. AngularJS is used for user friendly front-end and Express & nodeJS for backend. Analyzed api calls and ways to reduce them. Google APIs are used get route and cities on route as well as OpenWeatherMap api is used to gather weather information. Developed RestApi to get stored info of route & weather.

Angular, NodeJS, Express, Javascript,RestApi, Google Map API, OpenWeatherApi

Tom & Jerry — *Reinforcement Learning*

Trained agent Tom to catch Jerry in the game. Reinforcement learning is performed over grid-world environment. DQN and Deep Learning used to train model.

Python, Tensorflow, Deep Learning

EDUCATION

University at Buffalo, SUNY
M.S in Computer Science
Dec 2019   CPA - **3.62**

Walchand College of Engg, India
B.Tech in Information Technology
May 2015   GPA - 7.68/10

SKILLS

Programming: Python, Java, C++, C

Web: JavaScript, NodeJS, Django, Bootstrap, PHP HTML, CSS, RestApi

Database: MSSQL, MySQL, PostgreSQL, MongoDB, Hive, SQL

Big Data: Hadoop, Spark, Kafka, Hive, HDFS, Oozie, Sqoop, Data Analytics, flume

Cloud: AWS, SQS, SNS, EC2, S3

Other: Docker, VM, Git, Jenkins, Maven, Jira, Spring, JSON

Learning: Angular 7, Cassandra, HBase, GraphQL

OS: Linux, Windows

SDLC: Agile, Scrum

PATENT & AWARDS

- **Patent granted.** *US10250588*
  System and Methods for Determining Reputation of Digital Certificates Signer

- 2 Awards : **Passion & Focus** for excellent work done on multiple projects at Symantec

- Accolade: **Openness & Collaboration** for partnering with multiple teams at Symantec.

Course

- Algorithms
- Distributed Systems
- Data Intensive Programming
- Data Structures
- Databases
- Machine Learning
- Software Engineering

Hobbies