BIJAY KUMAR PATHAK

(503)-547-4801   [linkedin]   pathakbijaykumar@gmail.com
                                                                                                 Portland, OR

ACADEMICS

Michigan Technological University   **2012-2014**
M.S in Computer Engineering, Specialization: Machine Learning   **GPA: 3.53/4.0**
Coursework includes Data Mining, Advance Artificial Intelligence, Computational Intelligence, GPU and Multicore programming, Numerical Optimization, Mathematical Statistics, Time Series Analysis, Numerical Linear Algebra, Statistical Inference.

Tribhuvan University, Nepal   **2007-2011**
B.E in Computer Engineering
Coursework: Programming Languages, Algorithms and Data Structures, Data Mining, Operating System, Database, Software Engineering.

WORK EXPERIENCE

McAfee Corporation.   **April 2017 - Present**
Software Engineer   **Hillsboro, OR**
1. Currently, working on the APIs and backend service for Identity and Access Management service for all the products on Cloud.
2. Worked on microservices to process, enrich and parse events data from Kafka stream and AWS DynamoDB as storage.
3. Developed a custom Sqoop – like a tool with byte serialization to transfer Terabytes of data from relational databases (MySQL, MS SQL) to HBase using Spark API for Scala.
4. Devised automated deployment scripts for infrastructure/cluster management in AWS using Terraform and Amazon CLI.
5. Worked on the containerizing application and automated deployment using Docker, Kubernetes, and Terraform.

Intel Security (McAfee), Intel Corporation.   **August 2016 - Mar 2017**
Software Engineer   **Hillsboro, OR**
1. Worked on building a distributed Big Data Platform for Intel Security ePolicy Orchestrator with HBase as the NoSQL Datastore.
2. Worked on a different part of projects like custom reporting dashboard, custom query translator and continuous integration and continuous delivery.
3. Setting multi-node Hadoop Ecosystem cluster, configuration and fine tuning including Spark and Kafka using Puppet to automate.

Nike.   **Jan 2016 - July 2016**
Software Engineer   **Beaverton, OR**
4. Worked on Nike Next Generation Analytics Platform.
5. Worked with Apache Spark, Apache Airflow, AWS services including EMR, S3 and Apache Hadoop stack.

Cloudwick, Inc.   **2015 - Dec 2015**
Software Engineer, Intern   **Newark, CA**
1. Wrote custom Hadoop MapReduce to analyze the clickstream data to reduce the bouncing rate and improve conversion.
2. Worked with the client to move the data processing pipeline to Spark on YARN cluster from the RDBMS system.
3. Setting multi-node Hadoop Ecosystem cluster, configuration and fine tuning including Spark and Kafka using Puppet to automate.
4. Worked on building Nike Next Generation Analytics Platform using Apache Spark with Scala, Python and Apache Hadoop Ecosystem.
5. Worked on building analytics platform to consolidate all the analytics requirement using AWS services for batch, streaming jobs, ad hoc queries, and Data as a Service to end user.

Dept. of ECE, Michigan Tech   **Sept 2012 - Aug 2014**
Research/Teaching Assistant   **Houghton, MI**
Worked with the real-world and synthetic datasets to find numbers of triangle and clustering coefficient using the Eigenvalue Method.
Focused on increasing the speedups while maintaining more than 95% accuracy than the traditional method using GPU.

CERTIFICATIONS

Databricks Certified Developer on Apache Spark
Cloudera Certified Administrator for Apache Hadoop (CCAH)
Cloudera Certified Developer for Apache Hadoop (CCDH)
Machine Learning, Practical Machine Learning, Computing for Data Analysis from Coursera; Scalable Machine Learning EdX

TECHNICAL SKILLS

Language: Currently working (Java, Python, Node.js), during research work(C/C++, CUDA C), some experience (Scala, Go)
Hadoop Ecosystem: HDFS, MapReduce, YARN, Hive, Zookeeper.
Apache Spark: Spark, Spark SQL, Spark DataFrame, Spark Streaming
Apache Kafka: Experience in integration of Kafka to create real-time data ingestion pipeline and writing custom producer and consumer.
Docker: Experience in automation deployment using Docker, Kubernetes, and Terraform.
Linux: Experience in Linux Administration, Configuration, and File Systems.