Vivek

Address	Clarendon Rise, London	**MyBlogs**	http://www.machineintellegence.com
Phone	+44 73-841-65-841	**LinkedIn**	https://www.linkedin.com/in/vivekchaturvedi1/
E-mail	vivek.reeves@gmail.com	**My Code**	https://github.com/vivekreeves/
		Stack:	https://stackoverflow.com/users/3761442/reeves/

Lead Data Engineer

Lead Data Engineering with 11+ years of experience in handling big data and analytics projects. Overall technical lead/manager responsibilities including road maps, leadership, planning, technical innovation, development, performance. Provide technical and process leadership for projects, defining and documenting information integrations between systems and aligning projects. Involved in technical designing of the big data, analytics project and the implementation of the solution with the Spark/ML/Cloud technologies.

Responsibilities

- **Technical Lead:** Mentoring, Training , task allocation, Code reviews, Bug Fixing, Development, Deployment and SL3 supports.
- **Manager:** Project Planning, Costing and Estimation.
- **Architect:** Architecture and Planning, Project migration impact and analysis, RFP and POCs. Code migration on cloud and cost estimation.

Profile

- Proficient in Coding with **Python**, C#.
- Experienced in Data Engineering and building **Spark based framework.(Yaml,Sql,Json).**
- **Hands on with improving performance of Spark jobs.**
- Experienced in Machine Learning and Analytics using SK Lean, Tensorflow and Pytorch.
- Proficient in BI, Big Data problems (Hadoop), Hands on experience on **Hive,Impala, Sqoop, Spark, Spark-MLib. Spark-SQL.**
- Hands on with Cloud GCP- DataProc, AWS -EMR.
- Handled multiple projects using **Agile** -SCRUM,KANBAN and Activity mapping strategies.
- Proficient in **KDD, SEMMA and CRISP-DM** frameworks.
- Developed **data lineage** and **reconciliation** framework for Big data projects.
- Experienced in development, Migration and support of Data Analytics projects.

Career Contours

Organization	Designation	Role	Duration
Deutsche Bank-UK	Lead Data Engineer	Data Engineer	Jan 2020 till date
HCL technologies-India	Tech- Lead	Data Scientist	Jul 2017 to Dec 2019
Tata Consultancy Services	IT Analyst	Technical Lead	Oct 2015 to Jul 2017
Birlasoft (India) Limited.	Sr.Software Engineer	Technical Lead	Aug-2012 to Sep-2015
Ebix Asia SEZ Pvt. Ltd.	System Analyst	Programmer	Nov-2010 to Aug-2012

Servitium IT Solution Pvt.Ltd	Software Engineer	Developer	Oct-2009 to Sep-2010

Domain Experiences

Telecommunication with Nokia
Banking and financial services with Nationwide building society(UK), Deutsche Bank
Finance with General Electric(GE)
Insurance with WMIC
Gaming with Microsoft(Xbox GTSC)

Education

Course	College	University	Year of Passing
M.Tech	BITS Pilani	BITS	Pursuing
PG-Diploma	CDAC Noida	GGSIPU	2009
Bachelor in Technology (CSE)	JSS Noida	UPTU	2007
Diploma (3 year) in (M.CAD)	M.P.Polytechnic	BTEUP	2004
12th	Army School	CBSE	2001
10th	Army School	CBSE	1999

Technical Skill

Technologies	Big Data, Cloud- GCP- Dataproc, AWS-EMR, DataScience, Machine Learning, Deep Learning
Languages	**Python**,Java, C#, Scala.
RDBMS	Sql Server, Oracle.
Python Libraries	PySpark, NumPy, Pandas, MatplotLib.SkLearn,
Web Technology	ASP.Net ,Django, Flask
Deep Learning libraries	Tensorflow, Keras, NNabla
Release Management Tool	Visual Source Safe, TFS, StarTeam, SVN, **Git**, Bit Bucket
Tools	**Control-m,** Nifi, **Unravel**.
Big Data	Hadoop, HDFS, Map Reduce, Hive, Sqoop, flume, Kibana, Grafana, Spark-Mlib, **Spark-Sql, YARN.**
IDE	**Anaconda**, **Pycharm**, Visual- Studio, **Eclipse**.
Machine Learning	Tensorflow, scikit-Learn. NLTK,Keras,Pytorch.

Data Engineering

CDH	CDP	Cloud	Tools
Hive	Beeline	GCP- Data Proc.	Control-m
Spark	Spark	Azure-Databricks	Unravel
Sqoop	Sqoop	Amazon-EMR	AWS EMR, Pipelines
Impala	Impala		
Python	Python		

Machine learning Models

Supervised	Unsupervised	Computer Vision	Deep Learning
Regression	K-Means	GMM	LSTM-Seq-to-Seq
Logistic regression	NER	Haar cascade	Encoder Decoder
Decision Trees	Summarization	CNN	

Key Project Handled

Projects   **: Jan 2020 to till date**
Domain   **: Financial Domain**
Project   **: Program Reporting.**
Technologies   **: CDH5, CDP7, Spark2.4, Python-3.6**

Organization   : Deutsche Bank   **Tenure:** Jan 2020-till Date
Designation   **: Lead Data Engineer**
Roles & Responsibilities   **: Technical Lead (Spark)**

- Leading **Investment banking** Development project using **Agile scrum** mythology.
- Involved in **Designing** & **Development** Strategy.
- Develop more than 200+ transformation and helped in writing and integrating more than 1k transformation in Spark
- Tune the spark transformation by analysing DAG and YARN logs.
- Improved spark performance by 25%.
- Helped in mentoring and building the team of 30+.
- Designed the CICD pipelines and Agile setup for faster development and deployments.
- Increased the code commit velocity to 15% by improving the documentation and applying resource mapping.
- Manage stakeholder's expectations, escalations along with relationship building.
- Manage L3 support for the Production Support team to run the show.
- Maintain & Design high availability environment and also drive **DR exercise.**
- Involved in Migration and upgradation of infrastructure and cloud migration.
- **Capacity Planning and Housekeeping planning** for Cluster for smooth production run.
- Taking care of resource allocation, resource utilization and **Sprint Planning**.
- Created reconciliation tools that saved 15% of the work unit.
- Reconciled complex data using hash Strategy.

Project Name: Text Summarization, Clustering of KPI, Jira Automation.
Domain: Telecom
Technologies: Python, OpenCV, Tensorflow, Deep learning models, keras, Spark Mlib, ELK, SK Learn, Tensorflow.
Role: Data Scientist

- Leading **Telecom Industry** project using **Agile scrum** mythology.
- Involved in **Designing** & **Development** Strategy.
- Developed more than 15+ ML models and helped in writing and tuning more than 20+ models.
- Tune the pipelines and models by analysing DAG, Statical analysis.
- Helped in mentoring and building the team of 20+ developers.
- Helped in creating a ML framework that can clean the data and predict the value based on type of Data.
- Implemented Regression, classification , clustering , Decision tree , NER model on text data.
- Implemented and integrated Tensorboard on UI to implement PCA and Tsne plots.

Extra-curricular Activities and Certification

1. Achievements

a. Winner of Hackathon conducted by **IBM** for designing a tool **"Chota rasta" for predicting the minimum distance for cabs.**
b-Received the award for *"**Best Synergy Team***"* of the Quarter in Birlasoft while working as a tech- lead.
c-Received **BrandonHall Gold Award** for Best Use of Games for Learning in 2010.

Certifications	Institutions
Bash for Programmers	Educative
Python	Rice University
Investment Banking	Udemy
Machine Learning with Python	IBM
Scalable Machine Learning on Big Data using Apache Spark	IBM
Deep learning & Neural network with Keras	IBM

Personal Details

Name: Vivek Chaturvedi

Location: London, UK

Passport No: XXXXX-XX

Visa Status: ICT Tier 2. Valid till 2025.