UTKARSH KUMAR SINGH

3315 E Longview Avenue, Bloomington, IN 47408

utksingh@iu.edu | Tel: (812) 955-1508 | www.linkedin.com/in/u-singh

EDUCATION

Indiana University (IU), Bloomington, IN, United States *Aug 2016 - Present*
Masters of Science in Data Science   GPA: 3.7/4.0

Relevant Coursework: Data Mining, Advanced Database Concepts(SQL), SQL & NoSQL,
Exploratory Data Analysis in R, Data Analysis & Modeling in SAS, Social Media Mining(Python)

National Institute of Technology (NIT), Durgapur, India *Aug 2012 - May 2016*
Bachelor of Technology, Information Technology   GPA: 8.14/10

PROFESSIONAL EXPERIENCE

Westfield Gas & Electric, MA *Summer 2017*
Data Analyst Intern

- Worked on **Smart Meter Data Analysis** project whose aim was to analyze smart meter data and get important insights on electricity conservation measures and initiatives for utility as well as customers.
- Installed the latest stable version **Apache Hadoop-2.8.0** on the Oracle VM Virtual box and its other dependency softwares in the production environment.
- Completed batch ingestion of smart meter data into HDFS using a **flume agent on Apache Flume** 1.7.0.

National Informatics Centre (NIC), New Delhi, India *Summer 2015*
Summer Trainee

- Assisted Senior Database Administrator regarding the **configuration, security and monitoring performance** to resolve bottlenecks of 20 plus SQL Servers (2012).
- Implemented SQL best practices(proactive approaches) for setting jobs using **SQL server Agent** for database backups, Transaction-log growth montoring, Tempdb growth monitoring, monitoring Blocks & Deadlocks.
- Implemented & Documented **backup and recovery disaster plans** for optimum redundancy.
- Audited slow performing stored procedures & queries to optimize them for efficient performance.
- **Devised reports in SSRS** & **automated report creation** by deploying subscription & scheduled as per the end user requirements.
- **Generated and deployed SSIS packages** for migrating data between heterogeneous environments using various transformations.
- Strong points include excellent written and verbal skills, ability to work independently **with little supervision, goal orientated, strong desire to learn, able to manage multitasking and can work comfortably with other colleagues.**

Informatica Business Solutions Pvt. Ltd., Bengaluru, India *Summer 2014*
Marketplace Intern

- Worked on **PowerCenter,** the flagship ETL tool of Informatica & got acquainted with it.
- Developed Python scripts(for windows & Linux) to enhance the utilities of PowerCenter:

  Create relational connections **(Oracle, DB2, SQL Server, Sybase & Teradata)** in bulk to establish connectivity between databases & Informatica repository.

  Check validity of the relational connections after creating them.

TECHNICAL SKILLS

Databases: MYSQL, PostgreSQL, Oracle 11g express, SQL Server (2012, 2014, 2016) : [T-SQL, SSMS, **SSIS, SSRS, SSDT**]
Programming Languages: Python(Scikit, numpy, pandas), R, SQL, SAS, C, C++
Big Data Technology: Hadoop, Spark, Flume
ETL Tools: Informatica PowerCenter, SSIS
Data Visualization: Tableau, ggplot2 and ggvis (R), matplotlib (python)
Tools and Miscellaneous: R Studio, R Markdown, Latex, HTML, Microsoft Visual Studio

RELEVANT PROJECTS & COURSES

- **SQL Server Administration** Certification - Implemented Real world Production environment Administration techniques such as, **backup and restore procedures**, monitor health of SQL Server using **perfmon**, review storage, data and log file necessities, reorganize & rebuild table index, configuring security & setting up maintenance plans, Performance tuning using **SQL Profiler**, **Database Tuning Advisor** (**DTA**) & **DBCC scripts**, create & investigate issues using **Database Management views & extended events**, configure & monitor **High Availability Solutions** (**log shipping, database mirroring**, **replication** & **Always on availability** group technology), deploy **SSRS** for reports, deploy **SSIS** packages to migrate data.
- Created a predictive model to predict **Stock Market movements of 'Nike'** using supervised machine learning Principles to the tweets extracted from Twitter and developed a sentiment analyzer that can judge the type of sentiment present in the tweet. Analyzed the correlation between stock market movements and sentiments in tweets.
- Created a predictive model to accurately predict first booking destination for new Airbnb customers as part of the **Kaggle competition "Airbnb New User Bookings"** using advanced ML algorithms such as xgboost & secured a rank in top 20% of the competition.
- Secured a rank in top 20% of a "**Kaggle competition: House Prices Prediction**" while building a model to predict the prices of residential homes in Ames, Iowa, using feature engineering and advanced regression techniques.
- **Analytics & Visualization Challenge -** Performed data analysis & build visualizations using Tableau that uncovered significant trends in data and identified factors that could influence the sales and increase revenue of grocery stores across various demographics.
- Created a predictive model using SAS & R to predict whether an **employee will leave the company or not** by analyzing different numerical & ordinal attributes influencing an employees' decision.