

Khasimbabu Shaik

Software Engineer Lead

Software professional with 14 years of IT experience, have exposure to Design, Development, Testing and Maintenance of medium to large scale projects. I have work experience in middleware integration products. I am professionally trained and certified in Big Data and Data Science technologies. I am seeking a Big Data/Data Science analytics position to utilize my integration and analytical skills in the project development and growth of the organization.



khasimbabu.eai@gmail.com



(+91) 9985690604



Hyderabad, India



linkedin.com/Khasimbabu



github.com/Khasimbabu



Colab Notebooks/Khasimbabu

WORK EXPERIENCE

Genzeon Technology, Hyderabad — Software Engineer Lead Urban Outfitters, USA

MAY 2013 - PRESENT

I have been working in multiple roles in this Organization and thoroughly achieved my tasks and responsibilities in each role.

- Senior Software Engineer
- Staff Software Engineer
- Software Engineer Lead

Urban Outfitters

MAY 2013 - PRESENT

Urban Outfitters Inc., a lifestyle retail company, operates retail stores under the Urban, Anthro, Free People and Terrain brands. URBN maintains its business transactions by using different applications. Tibco acts as an Integration layer for data communication between these applications.

- Integrating the new applications, enhancing the existing applications development. Leading the team to achieve their deliverables on time.
- Involve in daily development, production monitoring and maintenance activities. Resolving the production issues and bug fixing
- Developed the Rest API, Database, WebServices and Messaging services using Tibco for integrating different applications (like Sterling, WMS, ACI, PXP, Vertex, Salesforce, GoogleCloud, Amazon, Ryder, Experian, Tmall, Mirakl, SessionM, Tradestone..)
- GitHub/SVN for code check in & Jira for tracking the Task updates, build plans and documentation. Slack for official communications.
- Has working experience on BigData Kibana Elasticsearch and Kafka messaging system.

Mphasis, Chennai — Delivery Software Engineer Department for Work and Pensions, UK

MAY 2006 - JAN 2012

DWP SOA, BLACC

SEPT 2009 - JAN 2012

DWP SOA and DWP BLACC are LegaSuite projects in which we automate already existing services in different benefit systems(JSAPS, ISCS, PSCS and DLACS) of DWP using Legasuite products.

TECHNICAL SKILLS

Middleware Technology

Tibco, LegaSuite

Programming Languages

Python, Java, Scala, C, C++ and Cobol

Databases

SQL Server, Oracle and DB2

Big Data

HDFS, YARN, MapReduce, Pig, Hive, Scala, Spark SQL, Spark Streaming, MLlib, MongoDB, HBase, Kibana Elasticsearch, Sqoop and Flume

Data Science

Python, Numpy, Panda, Sklearn, Matplot, Seaborn, NLTK

Messaging Systems

EMS, RV and Kafka

Networking

Restful Web Services, SOAP, HTTPS and WSDL

Miscellaneous

Cloudera, Cloudx, Postman, Jira, Slack, PagerDuty, GitHub

GENERAL SKILLS

Problem Solving

Creative Thinking

Interpersonal Skills

Leadership

Teamwork

DWP CSA, AMS

OCT 2007 - AUG 2009

The DWP CSA is all about Child Support, DWP gives pensions for Child support. TIBCO will act as a middle layer to collect and pass the Information between different applications in the project. Applications Maintenance and Support (AMS) team is responsible for the Maintenance & Support of DWP-Child Support applications.

MEDPRO RENEWAL SYSTEM

MAY 2006 - SEPT 2007

The Medical Protective company is the oldest professional liability Insurers in the United States. The main objective is to integrate the data between AS400 and Oracle systems. The Business Logic was built in TIBCO.

EDUCATION

NCET - JNTU, Hyderabad — *Bachelor of Technology*

- I have completed a Data Science Program from GUVI (IIT-M Advanced Programming & Master Data Science)
- I completed the Training program from APSSDC and Certified (Data Science and AI using Python)
- I participated in the Microsoft AI Classroom Series. I have completed the Microsoft Ignite Azure Data Scientist Challenge
- I completed the BIGData program from UpX Academy and Certified as a BigData developer

SELECTED PROJECTS IN DATA SCIENCE

◆ Item Demand Forecasting

- Dataset has Historical sales data with Item, Store and Sales details provided at the daily level for a period spanning 4 years
- Predicted the demand for the next 3 months at the item level (i.e., all the stores combined)
- Developed a model that predicts future demand and gives the best profit to the manufacturer (i.e., strikes a balance between the opportunity cost and retention cost)
- [Colab link](#) to project

◆ NLP – Sentiment Analysis of IMDB reviews

- IMDB Dataset has review text and sentiment of each text. Comma separated values with 50000 observations
- Pre-process the data (Remove Punctuation, replace short words, making text lowercase, Remove stop words, Lemmatization)
- Plotting most frequent words from positive reviews using bar chart
- Build the model for the prediction
- [Colab link](#) to project

SELECTED PROJECTS IN BIG DATA

◆ Flights and Airports data analysis – using data flow language PIG

- Dataset comprised of Flights and Airports information of US Domestic flights operated by large air carriers
- Datasets has comma separated values data format with more than 30 columns and 1300 records data
- We analyzed the data and extract the Top visited destinations, List of Cancellations due to bad weather, Origins with highest departure delay,

AWARDS

Genzeon - SpotLight award for Support Excellence & Exceptional Contribution

Genzeon - SpotLight award for Client Focus & Diligence

Mphasis - 5 yrs completion Pillars appreciation

Mphasis - Value Appreciation Card

LANGUAGES

English

Hindi

Telugu

max diversion routes information

- [GitHub link](#) to project

❖ **Retail Sports Store data analysis – using HIVE**

- There are 2 Dataset. Transaction data file has Transaction information and Customer's data file has Customer information
- Comma separated values with 50000 observations
- Created different tables and join them then grouping the parameters and found the amount spend by different age group people; it will be helpful for the business growth
- [GitHub link](#) to project

❖ **IPL data analysis – using SPARK SQL**

- The dataset contains IPL data. This data can be used to analyze some key statistics of players, teams, and match venues
- Data format is '|' separated values. There are 5800 observations
- We convert the file into Data Frame and derive some key statistics from the data by using Spark SQL
- [GitHub link](#) to project

❖ blogger.com/BigDatablog/Khasimbabu