Gus Segura

Principal Data Science Engineer, Lead: Big Data Architecture, Hadoop HDP, Oracle, Hive, Pig, NiFi, Sqoop, Python, Spark, Scala, Java, Data Pipeline, Kafka, Storm, Apex, Redis, AWS, EC2, EMR, S3, Redshift, Data Warehouse, Data Architect, SQL, ETL, Docker, Kubernetes, Git, Jira, Atlassian, Tableau, MicrosStrategy [Sacramento, Seattle, Remote].

Placerville, CA - Email me on Indeed: indeed.com/r/Gus-Segura/d682d81a6d06e851

Principal Data Science Engineer, Lead, Architect, Data Model, DevOps, Big Data, Hadoop, Oracle, Teradata, SQL Server, Hive, Pig, NiFi, Sqoop, Python, Spark, Scala, Java, Data Pipeline, Kafka, Storm, Apex, Redis, AWS, EC2, EMR, S3, Redshift, Data Warehouse, Data Architect, SQL, ETL, Git, Jira, Atlassian, Data Mining, Machine Learning, Tech: Python (TensorFlow, Theano), R, SAS, H2o, MR, C, C#, Bash, (csm,psm), Tableau, Splunk, MicroStrategy, OBIEE, Dashboards, Reports, Visualizations. Microservices, Kubernetes, Mesosphere, Git, Atlassian. Mongo, Couchbase, Redis, [Sacramento, Seattle, Remote].
[Please No CV/Resume Spammers].
Authorized to work in the US for any employer

WORK EXPERIENCE

Principal Data Science Engineer, Lead

Blueskymetrics - Seattle, WA - 2012 to Present

Responsibilities - Principal Data Engineering: Professional Consultant.
Principal Data Engineer, Lead: Design, development, and implementation of big data analytics platforms. Combining on premise and cloud technologies using data driven and streaming real-time analytics for data mining, sentiment analytics, discovery, customer behavior, customer experience, insight, metrics, incentives, promotions, retention and ad-hoc big data analytics.

Accomplishments
Develop data driven solutions utilizing current and next generation technologies to meet evolving business needs. Delivering solutions to business critical problems by interfacing with customers, leadership, stakeholders, resources and business managers. Develop real-time data ingestion and stream-analytic solutions for: Manufacturing: quality, yields, defect tracking, raw materials, vendor management, supply chain and logistics. Marketing: segmentation, propensity to purchase: combining sentiment analytics, customer behavior and profile to predict optimum campaign strategies. IOT Edge Analytics, Finance and Healthcare.

Specializing: architecture, data modeling, data pipeline, data quality, cleansing, performance tuining, devops, cluster management, data movement, etl, sql, tsql, nosql, map reduce, complex aggregation. Visualization: Tableau, MicroStrategy, JS, D3.js, HTML5, CSS, Docker for Realtime dashboards, reporting and key metrics connecting to rest api, hadoop, mongodb, oracle, ms sql server, mysql, redis. Data Mining: process mining, classification, clustering, regression, prediction, text data mining and sentiment.

Skills Used
Hadoop, Oracle, Hive, Pig, NiFi, Sqoop, Python, Spark, Scala, Java, Data Pipeline, Kafka, Storm, Apex, Redis, AWS, EC2, EMR, S3, Redshift, Data Warehouse, Data Architect, SQL, ETL, Git, Jira, Atlassian, Machine Learning: Python, R, Java, JavaScript, Node, MEAN, D3, Angular, R,R-Studio, H20, Jupyter, Zepplin, OBIEE, Excel, PowerBI and Tableau.

Principal Data Engineer - DEIII+

Amazon - Seattle, WA - 2010 to 2012

Principal Data Engineer, Lead in system design, experiment design, database model, etl, user data access rights, sql / nosql performance, custom metrics, data quality, user admin and ops support. EDW, Big data implemented in AWS with Hadoop, EMR, DynamoDB, Hive, Pig, Oozie, Sqoop, Oracle, MongoDB, Redshift, Redis, Kafka and storm.

Accomplishments
Customer behavior correlated to purchase, vendor, shipments, inventory and costs. Lowering cost by shipment, inventory and logistic optimization. Increase sales lift through relevance, recommendation and incentives via clustering, classification and segmentation. ACS: A big data machine learning relevance (classification) system used to improve the customer experience. ADQ: Data anomaly in time series system used to statistically analyze patterns in data quality.

Specializing: architecture, data modeling, data quality, cleansing, performance, devops, cluster management, data movement, etl, sql, tsql, nosql, map reduce, complex aggregation. Visualization Development: Java, JS, D3.js, HTML5, CSS for Realtime dashboards, reporting and key metrics connecting to rest, hadoop, mongodb, oracle, ms sql server, mysql, redis. Data Mining: process mining, classification, clustering, regression, prediction, text data mining and sentiment.

Skills Used
AWS, DDB, RDS, MS SQL Server, Postgresql, Oracle, Hadoop, Hive, HQL, Pig, Spark, Spark-SQL, Sqoop, Map Reduce, MongoDB, Redis, Spark-SQL, ETL, UDF, UDAF, Java, Python, R, SAS, SQL, TSQL, JavaScript, Node.js, D3.js, Angular.js, JQuery.js, Visualize.js, HTML5, CSS, JS, Graph.js, JSHint, JSLint, CC, R,R-Studio, SAS 9.4, OBIEE, ESSBASE, MicroStrategy, Excel and Tableau.

Principal Engineer

TADR - Google - Placerville, CA - 2008 to 2010

Responsibilities
Designed, developed and implemented a marketing segmentation system - A big data machine learning relevance (classification) system used to improve customer experience for Advertising with small business (Google AdWords). The system was developed as plug-in for Google Analytics allowing customers to match Ads more closely to they're preferred customer segment.

Accomplishments
Agile: Design, development, implementation: Enterprise Data Warehouse, Big Data and Oracle. Apache core Hadoop, Hive, Scala, Pig, Oozie, MongoDB, Java, Python, SAS 9, R, R-Studio, RDD, ETL, Pentaho 5.1, workflows and streaming with map reduce. Two week sprints, agile, scrum.

Specializing: architecture, data modeling, data quality, dba, cleansing, performance tuining, devops, cluster management, data movement, etl, sql, tsql, nosql, map reduce, complex aggregation.

Skills Used
Oracle, MS SQL Server, SQL, T-SQL, DBA, Java, JavaScript, Node, MEAN, Docker, Python, Matlab, Application, Apps, KVP, MR, MSTR, Aggregators, ETL, Data Modeling, Reporting, Excel, Tableau, Hadoop, Eclipse, Open Source Analytics, Big Data, Hadoop, Map Reduce, MySQL, Oracle, Graph.js, OBIEE, SSIS/ SSRS, HTML, CSS.

Principal Application Architect

Computershare Gmbh. - München - 2007 to 2008

Responsibilities
Architected, designed, developed and implemented FinVal™. An esop, equity, option financial analytic business intelligence system for Atiken Private Bank [In Deutschland]. This PaaS solution was developed to allow Computershare clients to expose the interface for position valuation and portfolio management (I.e. Allianz insurance employee esop valuations portal.) Group Leader: Managing DBA's, engineers and valuation analyst.

Accomplishments
Financial analytics: fair value of complex derivatives, trade feeds, equities, foreign exchange rates, forecasting, trend analysis, Black & Scholes, time value of money and statistical analysis. Design, development and implementation of valuation system.

Specializing: time value of money, prediction, sql, ssis, ssrs, data modeling, complex aggregation, etl. Visualization Development: Java, WebServices, HTML, CSS for dashboards, reports amd key metrics connecting to ws, oracle, mssql and others.

Skills Used
Agile, Scrum, Java, JavaScript, MSVC, C#, SSIS, SSRS, C/C++, Oracle, MS-SQL Server, MySQL, ETL, SQL, PL/SQL, TSQL, Oracle DBA, SQL Server DBA, Developer, Data Modeling, OLAP, OLTP, Statistics, Predictive Analytics, Algorithms, HTML, CSS, Java, JavaScript.

Principal Data Engineer
Google - Mountain View, CA - 2006 to 2007

Responsibilities
Principal Engineer – Architecture: Human Resources workforce analytics, payroll planning, turnover (retention), cost per employee, incentive compensative management, and productivity management. Managing a medium sized team of engineers and hr cost analyst.

Accomplishments
Integrated data from disparate sources and supported complex incentive compensation plans with multiple variable analytics and data visualizations. Managed employee recruitment and retention strategies analytics. All analysis was correlated with other information to align corporate performance and productivity targets.

Specializing: process mining, regression analytics, prediction, sql, nosql, data modeling, map reduce, complex aggregation, etl, data quality. Visualization Development: Java, JavaBeans, HTML, CSS for dashboards and key metrics connecting to oracle, essbase, hyperion and others.

EDUCATION

Bachelor of Science (B.S.) [Equil. Rockwell Automation] in B.S.E.E
University of Nevada, Reno - Reno, NV

Business Law, Applied Statistics, Analytics
Folsom Lake College - Folsom, CA

SKILLS

Big Data, Oracle (RAC, ODM, BI), MS SQL Server, MongoDB, Hadoop, Redis, Spark, Spark-SQL, Hive, Pig, Sqoop, Flume, Kafka, Zookeeper, Yarn, Mesos, R, SAS, NoSql, KVC, Map Reduce, AWS, Microservices, Docker, Mesosphere, Enterprise Miner, Rapidminer, Autonomy, Orange, TextBlog, NLTK, TM, Neo4j, Cypher, Redshift, DynamoDB, EC2, S3, EMR, DDB, SWF, MS Azure, HDInsight, SQLSVR, MySql. Splunk, Elasticsearch, Software: Java, Python, C/C++, Obj-C, Cocoa, Swift, JavaScript, AngularJS, JQuery, GraphJS, HTML5, CSS, Web Services, Node.js, D3.js, JSON, BSON, REST, SQL, TSQL, PLSQL. Open Source Analytics, Statistics: SciPy, NumPy, Matplotlib, Weka, Stata, Octave, Matlab, TensorFlow, Theano, Cuda GPU, Visualization, Tableau, Excel, MicroStrategy, OBIEE, Access, PowerPivot. (10+ years)

LINKS

http://www.blueskymetrics.com

https://www.google.com/+BlueskymetricsSeattle

https://www.linkedin.com/in/gussegura

GROUPS

IEEE

TKE

FEMA NIMS Search & Rescue

ADDITIONAL INFORMATION

[No CV/Resume Spammers Please]. Part-Time, Remote is OK. Considering FTE Roles that include Leadership or Principal Level Architecture - Dev.