# PRASHANTH BALASUBRAMANI

Email: [prashanthbalasubramani@gmail.com](mailto:prashanthbalasubramani@gmail.com) github.com/Prash74 Phone: (812) 318-2839 linkedin.com/in/prashanthbalasubramani

**ACADEMIC QUALIFICATION**

**Indiana University**, Bloomington, Indiana Expected Graduation May 2017

Master of Science (Computer Science) GPA – 3.31 / 4.0

**Kongu Engineering College**, Erode, India May 2013

Bachelor of Engineering (Computer Science and Engineering) GPA - 8.12 / 10.0

# TECHNICAL SKILLS

* **Languages:** Java, Python, C/C++, JavaScript, Shell Scripting
* **Data Tools:** Hadoop, Hive, Ab Initio, Apache Zookeeper, Spark
* **Databases:** MySQL, Teradata, MongoDB, DB2
* **Miscellaneous:** Ansible, Agile methodology & Scrum Software Development, ControlM

# PROFESSIONAL EXPERIENCE

**Indiana University** **Bloomington** August 2015 - December 2016

**Associate Instructor (Python, Ansible, Shell Scripting, MongoDb)**

* Worked under Prof. Geoffrey C Fox on his Big Data and Analytics Course as an Associate Instructor
* Implemented Bench Marking tools to setup and run projects across different cloud platforms such as Chameleon, AWS and future Grid
* Developed a Ansible deployment scripts for configuring virtual machines on Enterprise cloud providers

**Capital One Financial Services** October 2013 - June 2015

**ETL Developer (Java, Hadoop, Shell Scripting, Ab Initio, Scrum)**

* Developed Hadoop MR and Spark scripts for transferring data from Relational Databases into Hadoop clusters
* Designed and supported real time data warehousing jobs for manipulation of Credit card data from external sources to Relational and Non-Relational Databases
* Implemented a Shell based Test data generation tool to speed up the testing process in our team. This allowed the testers to spend less time on data creation which helped in speeding up team releases
* Developed a simpler batch job design which reduced the number of standalone jobs needed for processing credit card data. This reduced the job execution time by 30% which resulted in more data processing
* Displayed exceptional skills on Agile and Test Driven Development methodologies

# ACADEMIC PROJECTS

* Designed a python based POS tagger using Bayesian Network implementation and Natural Language Processing. Using a tagged Data corpus as the training set, we trained our model to tag the parts of speech of every word in a sentence. Achieved a baseline accuracy of 95% **(Python)**
* Implemented a Web Search Engine using PageRank algorithm which fetches webpages based on Search key. and used Apache Hadoop to build inverted page index for web links stored in HBase. Designed MapReduce jobs to build inverted index tables from web links **(Java, Apache Hadoop, Apache HBase)**
* Designed a classifier for speech act classification on conversational data. Modeled the classifier using Support Vector Machine (SVM) on two major annotated dialogue act corpus - Switchboard Dialog Act Corpus (SwDA) on Telephonic Conversations and ICSI Meeting Recorder Dialog Act (MRDA) corpus on Meeting Conversations **(Python, NLTK)**
* Developed an in-memory UNIX based file system and file system operation on an ARM based processor. Also, implemented future mechanism for a simple chat client for transmitting TCP/IP and UDP packets across network **(C, Unix)**
* Built Machine Learning models such as Decision Tree, SVM, K-Nearest Neighbor, Adaboost, Bagging, Logistic Regression and XGBoost and compared their performance on credit card data to detect fraudulent transactions and credit card defaulters **(Python, Machine Learning)**