Mobile: +91-9706706659
Email: kashyapbarua@gmail.com

Kashyap Barua

Address: King's Palace 9

Internship

Summer Training	EMC^2 ACADEMIC ALLIANCE	Summer 2016

Big Data and Data Analytics
- Setting up of a Hadoop Environment on a Linux system. Installing various tools like PostgreSQL, Hive, Pig that are required in the Hadoop Environment
- Working on retail and census data for analytics. Some of the work involved cleaning up and preparing of data for analytics.
- Using R programming on RStudio for data visualization to make inferences and conclusions from the datasets.

Education

Bhubaneswar, Odisha	KIIT University	Fall 2014 – May 2018

- B.Tech. in Computer Science and Engineering CGPA: 7.02 (Upto 6th Semester)
- Undergraduate Coursework: Programming Tools and Techniques, Software Engineering, Object Oriented System Design, Operating System, Database Management System, Design and Analysis of Algorithms

Technical Experience

Projects

- **Flight Tracking System.** An implementation of Python language on a framework known as Jupyter for data visualizations and analytics of data. Worked on a flight tracking system using Intel's Galileo board and RTL-SDR (Automatic Dependent Surveillance Broadcast) transponder to generate flight data.

- **Call Record Analysis**. Implementation of a Linux based open source framework known as Bandicoot for the analysis of call record data generated. The call logs and texts were gathered after generation which were then analyzed for required conclusions and inferences.

- **Twitter Feed Analysis**. This projects involved the students gathering the twitter feeds from their respective accounts. The feeds were used to perform sentiment analysis using a Python based framework known as Jupyter.

- **Document Analysis.** An implementation of NLTK (Natural Language Processing Toolkit) for the extraction of entities from a given passage. This is followed by identification of relations between the extracted entities using POS (Parts-of-Speech) tagging. The final process involves the identification of topics being discussed in the passage.

- **Automated Answer-Sheet Evaluation System.** This system uses OCR (Optical Character Recognition) to detect the hand-written text in answer sheets (in English language) which is then converted to machine interpretable format. The converted texts are then processed using NLTK and machine learning methods to mark the candidate, based on keywords and test cases.

- **Cognition Online Test System.** This online test system consists of a web server with database facility. Students can connect to the server through internet to take tests which includes aptitude, technical and English tests. Examiners can also connect to the server to evaluate and setup up question papers. The online exam system offers Ranking reports, detailed result analysis and time management reports.

Publications

- **Trends in Big Data**
Accepted at CSI Communications (November 2016 issue)
Authors: Dr. Bhabani Shankar Prasad Mishra, Kashyap Barua

Languages and Technologies

- **Language and Libraries** C++; C; Python; Core Java; HTML; MySQL; CSS; Hadoop; R; PostgreSQL; Pig; Hive; Bootstrap; MapReduce
- **Software** R Studio; GitHub GUI (for Windows); Adobe Illustrator 2015; Adobe Photoshop CC 2015; FL Studio 12; Oracle 11g
- **Database** MySQL; Oracle; Greenplum DB

Additional Experience

- **Student Coordinator at ICDCIT** *(International Conference on Distributed Computing and Internet Technology),* **(Feb 2016 - Present)**

  Managed the publicity teams and the graphics team. Some of the tasks include designing of posters and contents for social media marketing. Analyzing of media coverage in social media websites.

- **Student Coordinator at FICTA** *(Frontiers of Intelligent Computing: Theory and Applications),* **(Sept 2016 - Present)**

  Managed the registration team as well as the sessions that were held as the part of the conference. Also, manage the various social media platforms of the conference like Facebook page and the official website of the same.

- **Student Coordinator at ICCIDM** *(International Conference on Computational Intelligence in Data Mining,)* **(Nov 2016 - Present)**

  Managed the registration team and the various sessions and paper presentations that were organized as a part of the conference.

- **Student Coordinator at EMC^2 Academic Alliance, (Mar 2016 - Present)**

  Managed the various training programs that are organized by EMC^2 Academic Alliance, KIIT. Also, complete various hands-on materials on Big Data and Cloud Computing that is provided as projects by the company. As the part of the summer program, I worked on various topics such as PostgreSQL with Greenplum database, R programming, Machine Learning concepts (Naïve Bayes classifier, Logistic Regression, Linear Regression), Hadoop, Pig and Hive.