**Email :** [haince.pen@rit.edu](mailto:haince.pen@rit.edu)

**Phone :** +1 585-358-1098

**Haince Denis Pen**

[www.linkedin.com/in/haincepen](http://www.linkedin.com/in/haincepen)

31 River Court, #2103, Jersey City, NJ 07310

**Objective:** Data Enthusiast with 4+ years of experience, seeking opportunities in Data Science, Data Engineering and Big Data Analytics.

**Work experience**:

* **Technology Analyst | Morgan Stanley | New York City, New York *August 2016 – Present***
* Currently a part of Morgan Stanley’s TAP program, working with their data and technology services division.
* **Data Scientist / Big Data Analytics Intern | Ericsson | San Jose, California *May 2016 – August 2016***
* Worked with SLI team to optimize and improve customer experience, increased retention, and network issue identification.
* Acting Project Manager with Pricing Analytics team to coordinate activities, provide technical expertise and manage 9 resources.
* **Graduate Research Assistant | Rochester Institute of Technology | Rochester, New York *August 2015 – May 2016***
* Research on designing a scalable data warehouse over big data, addressing the issues of data processing and user performance.
* Research on ontologies for migration of relational data warehouses to big data platform for greater scalability and ease.
* **Data Engineer Intern | HealthCrowd | San Mateo, California  *September 2015 – November 2015***
* Involved in data engineering/modeling using relational and NoSQL databases for the startup in healthcare domain.
* Built and tested the pipeline to migrate data from relational data warehouse to big data platform to tackle scalability issues.
* Performed data aggregation, analysis and research on optimization capabilities on campaign data and provided the KPIs.
* **Database Analyst | Rochester Institute of Technology | Rochester, New York *November 2014 – October 2015***
* Worked with the SPA department to and built the grant tracking infrastructure to manage funds in excess of $100 million.
* Developed dashboard for forms, reports and charts. Automated generation of Letters in MS word and Outlook using VB.
* Re-structured the existing queries and altered the reporting infrastructure for better performance.
* **Sales Analytics Intern | FedEx | Memphis, Tennessee *June 2015 – August 2015***
* Identified key business initiatives and strategic sales prospects for FedEx's SameDay City product by building dashboards.
* Automated almost 95% of manual reporting, improved time efficiency by 99%, integrated ability for data mining and BI.
* Developed reports to manage the sales revenue split, customer retention and daily shipment volume trend analysis.
* **Teaching Assistant | Rochester Institute of Technology | Rochester, New York**   ***August 2014 – May 2015***
* Worked with the faculty to assist in data modeling, design, management, and troubleshooting of databases in class.
* **Software Engineer | Atos | Mumbai, India**   ***July 2012 – July 2014***
* Performance and tuning in Oracle database using SQLT, AWR, ASH, ADDM reports, cloning, import-export and re-org activities.
* Performed business analysis; Involved in stakeholder meetings for design and deployment of the database architecture.

**Skills:**

* **Big Data:** Hadoop, Sqoop, MapReduce, Hive, Spark.
* **Databases:** Oracle, MySQL, MongoDB, Cassandra.
* **Data Mining | ETL:** Weka, SPSS, Excel, Pentaho spoon.
* **Data Visualization:** Tableau, ArcGIS, Google Charts.
* **Languages:** Python, Scala, Java, SQL, HiveQL.
* **Operating Systems:** Linux/Unix, Windows.
* **Tools:** Eclipse, PyChram, Intellij, MySQL Workbench, Cloudera Quickstart, Virtualbox, VMware.

**Education:**

* **M.S. Information Sciences and Technologies | Rochester Institute of Technology | New York *August 2016***

**Coursework:** Big data analytics, data warehousing, mining, data modeling, analysis, visualization and business intelligence. ***GPA:*** *3.70*

* **B.E. Information Technology | University of Mumbai | India**   ***July 2012***

**Coursework:** Data structures, algorithms, data warehousing, software engineering, artificial intelligence, project management.

**Research:**

* **Analyzing in-game intelligence in NCAA basketball players’ tweets *Spring 2016***
* Research on correlation analysis between players’ emotional state on outcome of the game, under guidance of [Dr. Raj Murthy](http://www.rajmurthy.com/rajmurthy/).
* **Big Multi-dimensional data marts for analyzing unstructured data *AY 2015- 2016***
* Research on design and implementation of a scalable data warehouse architecture over big data, under guidance of [Dr. Jai Kang](https://www.rit.edu/gccis/jai-kang).
* **Spammer Detection by extracting message parameters from spam emails *September 2013***
* Co-authored a research paper ‘[Spammer Detection by extracting Message Parameters from Spam Emails](http://research.ijcaonline.org/volume78/number10/pxc3891232.pdf)’ in IJCA Journal.

**Projects:**

* Presidential Polls - Big Data Analytics on unstructured data
* Predicting Titanic’s survivors - Data Mining on Kaggle dataset
* SameDay City Analytics/reporting Infrastructure (FedEx)
* Product Sales Analysis - Relational Data Warehouse
* Yelp Yourself - Information retrieval and recommendation
* Data Analysis using Regression - Statistics

**Leadership and honors**:

* Awarded Graduate Research Assistantship in Big Data domain, with funding and 45% tuition scholarship at RIT (AY 2015-16).
* Independent Coursework (20+ Courses) completed from top universities and companies through MapR, Coursera, edX etc.
* ‘Data lakes for Big Data’ certification completed from EMC. ITIL V3 Certification in service management completed with 90% score.