PUNEET SINHA

**Data Science and Consulting** at Infosys Ltd. ***Contact No****: +91-8888835462*

***Experience*** *: ~7yrs* ***E-Mail****:* [*puneet.punu.sinha@gmail.com*](mailto:puneet.punu.sinha@gmail.com)

***Current Location*** *: Pune* [*Puneetsinha@yahoo.com*](file:///C:\Users\puneet.sinha\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\6VFIGVOJ\Puneetsinha@yahoo.com)

***Github:-*** [*https://github.com/puneetsinha*](https://github.com/puneetsinha) ***DOB: 3 June 1989***

***Linkdin:-***[*in.linkedin.com/in/puneetsinhaDS*](file:///C:\Users\puneet.sinha\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\6VFIGVOJ\in.linkedin.com\in\puneetsinhaDS) ***Address***: *Asha Bhawan, SR No.280,Lohegaon*

***Kaggle:-*** [*https://www.kaggle.com/puneetsinha*](https://www.kaggle.com/puneetsinha) *Pune, Maharastra-47*

***Hackerrank***:- [*https://www.hackerrank.com/puneetsinha*](https://www.hackerrank.com/puneetsinha)

# **Career Objective**

Data Scientist with a passion for learning and implementing new concepts. Currently working on designing and developing data driven features for our product. Working on predictive analytics use cases for Customer Success Management like Churn prediction, up-sell/cross-sell models, health scores, Customer Segmentation etc.

Motivated by the Idea of technology transforming society and particularly interested in using Big Data

Analytic to solve real life problems. Evangelizing data science in my free time.

# **Experience Details**

|  |  |
| --- | --- |
| **Infosys Ltd** | |
| Designation and Tenure | Data Scientist || Nov 2016 till present |
| **Vodafone Shared Services** | |
| Designation and Tenure | Data Analyst (Payroll: UST-Global) || Oct 2014-May 2015  Data Scientist (Payroll: Vodafone) May 2015-Nov 2015  Data Scientist (Payroll : IBM)|| Nov 2015 – Oct 2016 |
| **Amdocs Development Centre India** | |
| Designation and Tenure | Senior Subject Matter Expert || Dec 2010-Aug 2014 |

**Technical Skills**

|  |  |
| --- | --- |
| **Big Data tools** | **PIG,HIVE,SPARK, CDH 5.8,HDFS 2.7** |
| **Data analysis Tools** | IPython notebook, Rstudio, Excel, Spyder ,Pycharm |
| **Data processing** | R/Python/Hive/SQL/Pig/Excel |
| **Programming Languages** | Python,R |
| **Databases** | PL/SQL(Toad, SQL Developer), Mysql,DB2 |
| **Scripting Languages** | Shell Scripting, SED & AWK, |
| **Python& R libraries(major)** | Pyspark,Sklearn,Pandas,scipy,matplotlib,seaborn,numpy, [e1071](http://cran.r-project.org/web/packages/e1071/)**,** [rpart](http://cran.r-project.org/web/packages/rpart/)**,** [nnet](http://cran.r-project.org/web/packages/nnet/)**,** [randomForest](http://cran.r-project.org/web/packages/randomForest/)**,** [caret](http://cran.r-project.org/web/packages/caret/)**,** [glmnet](http://cran.r-project.org/web/packages/glmnet/)**,** [gbm](http://cran.r-project.org/web/packages/gbm/)**,** [tree](http://cran.r-project.org/web/packages/tree/), etc |

# **Educational details**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Qualification | %age | Grade | School/college | Board |
| HSC March’06 | 72% | Distinction | KENDRIYA VIDYALAYA NO.1 JAMNAGAR | C.B.S.E |
| SSC March’04 | 74.4% | Distinction | KENDRIYA VIDYALAYA NO.1 JAMNAGAR | C.B.S.E |
| BACHELOR OF ENGINEERING (Computer Science) | 71% | Distinction | C.K.PITHAWALA INSTITUTE OF TECHNOLOGY | VNSGU |

# **Experience**

* Sound understanding of below statistical techniques **Logistic/Linear/Lasso/Ridge Regressions, Segmentation, Decision trees (Bagging and boosting), Recommendation Engines, Support Vector Machines, Dimensionality Reduction(PCA, feature selection), Neural Networks, NLP**
* **Worked on Optical Character recognition (Tesseract-OCR) to make out text out of image**
* **Data Visualization (matplotlib, ggplot), Text Analytics (BNS, Countvectorizer, TFIDF), Topic Modelling**
* Ability to interpret results of above mentioned Statistical techniques
* Very good knowledge of Big Data Technologies like ***MapReduce/Hadoop /HIVE/PIG/SPARK***
* Ability to translate Business Problem to a Statistical Problem and Statistical solutions to a viable Business solution

Always have worked in guidance of managers and upper management as always secured ***high ratings from management of AMDOCS, Vodafone and IBM.***

* **Onsite Experience of Sao Paulo Brazil (3 months)** for requirement and business analysis.
* **Onsite Experience of Tal Aviv Israel (3 months)** for requirement and business analysis.
* **Onsite Experience of Jakarta, Indonesia (3 months) f**or requirement gathering and analysis.
* **Automation using Shell scripting/Python scripting** (Real time scripts which analyze customer data and works like a backbone in analysis)
* **Web-Services Testing, Middleware testing, TIBCO**, Strong knowledge of **Telecom Billing concepts** – *Mediation, Rating, Billing Specially in* ***AMDOCS Ensemble and Enabler 7.5 and 8.1****,*

**Project Details Current**

|  |  |
| --- | --- |
| **Title** | **Develop End to end solution to classify mails for BPOST** |
| **Client** | **BPOST Belgium** |
| **Role** | **Lead Data Scientist** |
| **Technology** | **Python for data Science, Tesseract OCR,Translation API** |
| **Database** | **TOAD and SQL developer for oracle** |
| **Project Summary** | bpost has a special reduced tariff program for direct mail, meant to motivate companies to keep spending their marketing budget on the paper medium and boost DM volumes. However, customers have always tried to also send regular letters under DM tariffs as it is cheaper   * Extract text content from DM samples. The input will be PDF files which will be converted to text by the application developed by the Supplier. * Perform Natural Language Processing on top of the extracted text to determine the applicability of DM tariff category. * Provide analysis and insights on the data and proactively asking for any data missing that can have impact on improving accuracy * Provide exploratory data analysis report and insights like variable importance, outliers or any anomaly, pattern discovery, hypothesis testing report, relationship among variables, etc. * Provide statistical model and formulation of problem, Predictive model output on train and test, and Performance measures etc. Provide basic user interface through which users can run the application to provide a classification for new PDFs |

|  |  |
| --- | --- |
| **Title** | **Increasing Mobile Operator Revenue Out of VAS**. |
| **Client** | **Vodafone** |
| **Role** | **Data Scientist** |
| **Team Size** | **20** |
| **Project Summary** | * Providing Analytics solutions, includes, developing predictive models, requirement understanding, data preparation, exploratory data analysis, developing machine learning models, models validation and final model delivery * Providing consultancy to the clients for their business problems and effectively implementing the best analytical solution * Presentation to the customers on our capabilities in Analytics   Requirement gathering and assessment from customers, sales team, business leaders and subject matter experts   * Preparing Assessment documents, Design mock-ups/prototypes, screen pages, project delivery plan and power point presentations * Propensity model for up-sell/cross sell, deriving actionable insight and automating the data science workflow |
| **Results** | Able to up-sell, cross-sell VAS services to ~18 lakhs VAS customers.  Able to add ~7 lakh existing customers who took VAS services. |

|  |  |
| --- | --- |
| **Title** | **Churn Prediction Model** |
| **Client** | **Vodafone** |
| **Role** | **Data Scientist** |
| **Team Size** | **40** |
| **Project Summary** | * Feature Selection for the model, Maintaining the model with the new customer trend * Trying different models of ML to see the learning curves, Feature Engineering with ongoing data Analysis * Preparing Assessment documents, Design mock-ups/prototypes, screen pages, project delivery plan and power point presentations * Coordination with multiple teams and stakeholders to support product development across locations * Deep dive multivariate analysis on transactional & aggregate data * Design and Implementation of Predictive Analytics use cases for Customer Success Management. * Design and development of Data Science features for Customer Success Management. |
| **Results** | Decrease in Churn from 0.87% to 0.71% and making it stable |

|  |  |
| --- | --- |
| **Title** | **Vodafone Netherlands B2B Plan suggestion || NET Brazil B2B Revenue analysis** |
| **Client** | **Vodafone Netherlands** |
| **Role** | **SSME Business** |
| **Team Size** | **30** |
| **Project Summary** | Bill to Bill is a critical Implementation service aimed to ensure the following in case of billing system replacement:   * + - Invoicing correctness     - No leak of revenue     - No impact on end customers   This is achieved by comparing large volumes of Legacy system outputs, with the new system matching outputs:   * + - Rating results (unbilled usage)     - Bill charges at various levels     - Sample population of printed bills      Ongoing activities monitoring and handling - AMC, TRB (via TRB Gui), ongoing rejects handling.  EOD & EOM monitoring - Make sure EOD/EOM has finished successfullybefore starting EOC flow. During - Monitor the cycle run and make sure nothing is stuck in the middle After EOC has finished - handle rejects and send report to management for status , performance and quality. Workaround management - Create and manage work arounds - all workarounds should be approved by DBA and also checked from business aspect -In case there is a major impact on revenue - this requires approval by the business. Workarounds should be removed once the permanent fix has been initiated. |

**HIGHLIGHTS:**

* Felicitated with a total of 10 Appreciation Certificates from the clients for the outstanding contribution (for every Release worked on.
* Sent onsite 3 times to interact with customer for 3 months for requirements data analysis and report making
* Rewarded with 3000 US dollars for being single and effective contributor at AMDOCS Brazil

**Certifictions:**

Data science Specialization 















***I hereby declare that the above information is true to my knowledge.***

***Yours faithfully***

***PUNEET SINHA***