Umang Mahant

**Mobile:** +91-8329561960, +91-8149773957

**E-Mail:** [umang1611.annes@gmail.com](mailto:umang1611.annes@gmail.com)

**LinkedIn:** <https://www.linkedin.com/in/umang-m-2296bab5/>

**GIT** : <https://github.com/Umang1611>

A Data enthusiast with 3+ years of experience in Data Science, Analytics, consulting etc. supporting business domains like – Insurance, Health care, Sales. Involved in Python open-source community and passionate about deep reinforcement learning, targeting assignments as **Data Scientist/ Machine Learning Engineer**

**Target Location**: Mumbai, Pune, Hyderabad, Bangalore

**PROFILE SUMMARY**

* **An enthusiastic professional** with **nearing 4 years of experience** in implementing machine learning Techniques
* Strong understanding of the **Machine Learning lifecycle**, feature engineering, training, validation, scaling
* Engaging with **stakeholders** to produce clear, compelling, and actionable insights that influence product and service by creating interactive **dashboards** using **Tableau**
* Strong acumen in Data Exploration**, Data Treatment/ Processing** for efficient model performance and stability
* Worked with **Statistics** Methods-Hypothesis Testing, understanding **Distributions**, Sample tests, **ANOVA, Market Basket Analysis, RFM**
* Proven skills in **predictive modeling** using various Machine Learning Models **ANN, Classification, Regression, Clustering**
* **Managed a high-volume workload** within a deadline-driven environment and gained trust from clients and higher management
* Day to day activities included dealing with **Data Visualization** using **Tableau, Seaborn, Matplotlib**
* Developed and maintained **Live dashboard** with **ELK stack - Kibana**
* Strong experience on programming languages such as **Python, Visual Basic /VBScript** (Excel), **SQL**
* Excelled in gathering and understanding requirements of clients
* Organizing multiple meetings for various discussions such as **Feature selection** with the **Subject matter experts**
* Exposure in generating **KPIs,** selecting methods and techniques for obtaining results for business growth

|  |  |  |
| --- | --- | --- |
| **CORE COMPETENCIES** |  | |
| IT Project Delivery | Machine Learning Lifecycle | Statistical Analysis |
| Database and Technical Skills | Client Interactions | Data Visualization |
| Requirement Gathering | Problem Solving | Domain – Healthcare and Insurance |

**WORK EXPERIENCE-SUMMARY**

# Since Mar’20-Present as Data Analyst at Allianz, Pune Key Result Areas:

* Collaborate with a team of data scientists and engineers to embed analytics into the business decision processes.
* Pre-processing, Cleaning data along with statistical analysis and machine learning models to predict fraud claims
* Highlight findings to business stakeholders and contribute to weekly meetings with the management team

**Result:** Model predicted Fraudulent claims with an accuracy of 88% after solving imbalance data issue and Feature Selection

# Aug’17-Jul’19 as Associate Consultant II at Atos, Pune Key Result Areas:

* Customer Analysis and segmentation to target groups with customized product promotions.
* Identify and interpret trends or patterns in complex data sets, to recognise similar customer batches to enhance customer experience, prevent customer Churn resulting in effective business growth.
* Using RFM and clustering model outputs for micro-segmentation for detailed analyses to develop KPIs

**Result:** The machine model could classified churn customers with 82% accuracy and could generate a benefit of $4,319 with only 5,000 sample cases

# Highlights:

* Esteemed by **KUDOS International Award** in 2018 for outstanding performance at Atos Company
* Bagged **SPOT Award** for **outstanding performance during client visit** in 2018
* Received the **SPOT Award** in 2019 for being an active team player and **critical on-time delivery**
* Awarded with **ASHWA SPOT Award** for excellent performance as a new hire in 2020

|  |  |  |  |
| --- | --- | --- | --- |
| **IT SKILLS** |  | | |
| Statistics and Probability | Hypothesis testing | Data Visualization | Data Analysis |
| Market Basket Analysis | Machine Learning Models | Predictive Analysis | RFM Analysis |
| Ensemble models  Tableau Visual basic | Artificial Neural Network  Python VBScript | Clustering PCA  Oracle SQL Java | AdaBoost |

**INDUSTRIAL PROJECTS**

**Employer:** Atos, Pune

**Period:** Aug’17-Jul’19

# Project: Customer Analysis and Segmentation

**Tools:** RFM, unsupervised learning, Clustering (Partitioning, Hierarchical), K- means, Tableau

# Description:

* Project’s main goal was to truly get to know its customer base for **tailored customer relationship management** and identify sources of **growth by customized product promotion on homogeneous groups**.
* **Demographic, transactional and Health** Data of historical customers was extracted from multiple **SQL** tables. A Total of 2.3 lakh customer details with 17 features data was structured for Benefit Segmentation Analysis.
* Initially with **RFM matrix** customer bins could be interpreted and then **micro-segmentation** was done using mined data and cluster Partitioning techniques like **K-means**. Clusters were visualized with **PCA-** 91% variance was covered**.**
* **Key Performance indicators** were identified such as **Top claim categories** (monetary and volume) per cluster, Changes in **Coverages**, Rate of increase of **premiums** per cluster, **Churn** Probability per cluster

# Project: Customer Churn Prediction and Retention

**Tools:** liner Regression, Classification, Feature selection, EDA, SQL, Python, RFM, ANN

# Description:

* Data with a total of 4.5 lakh customer details with 19 features. Post Identifying opportunities for improvement in data cleansing, Pre-processing, EDA was performed to obtain few insights concerning churn customers.
* **Bootstrapping** was used to create 3 samples of Churn to Not churn ratio 80:20, 70:30, 50:50 with train-test as 70:30 each
* **Logistic regression** and ANN performed almost evenly well, but only logistic regression model provides insights in the variables which are important to predict customer churn and develop retention techniques

**Employer:** Allianz, Pune

**Period:** Aug’19-Present

**Project:** Detection of Fraudulent Insurance claims

**Techniques**: EDA, Data analysis, Ensemble models (Boosting, Bagging), Random Forest, AdaBoost, IHT, SMOTE

# Description:

* Working with structured data with 273 features and 4.92 lakh records with a proportion of 0.8% fraudulent claims.
* Contributed to **Data Processing** by treating missing values, outliers and imbalance using **SMOTE, up-sampling** since the proportion of data was a major issue for model performance.
* **Feature selection** was done by consultation with subject matter experts. 70-30% ratio for Train-Test Split
* Performed **Cross validation** on models like **Random Forest** resulted in 81% accuracy and **recall** of 0.68 which was enhanced using **IHT technique** and **AdaBoost** to **accuracy 88%** and **recall** of **0.81**

**ACADEMIC DETAILS**

* **Distinction** in **B.E. (Computer Engineering)** from **AISSMS IOIT**, University of Pune in **2014 - 2017**
* **First Class** in **12th** from **N. Wadia**, Pune, Maharashtra State Board in **2012 - 2013**
* **Distinction** in **10th** from **St. Anne’s Convent**, Pune, Maharashtra State Board in **2010 - 2011**