Zamir Siddiqui

Experienced Machine Learning Architect and NLP expert with around 15.3 years of experience in Banking, E-commerce, Pharmaceuticals & Healthcare, Telecom and HR domains. Excellent reputation for resolving problems, improving customer satisfaction and driving overall Data Science related Solution by using different ML/DL algorithms and tools. Consistently saved costs while increasing profits.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | |  | **Education** |  |  |  |  | | --- | --- | --- | | **2000-07 - 2003-06** |  | MCA: Computer Science and Applications  Aligarh Muslim University - Aligarh |  |  |  |  | | --- | --- | --- | | **1998-07 - 2000-06** |  | M.Sc.: Operations Research  Aligarh Muslim University - Aligarh |  |  |  |  | | --- | --- | --- | | **1995-07 - 1998-06** |  | B.Sc. (Hons): Physics  Aligarh Muslim University - Aligarh |  |  |  | | --- | --- | |  | **Work History** |  |  |  |  | | --- | --- | --- | | **2020-06–Current**  **2017-06 – 2020-05** |  | Data Science Architect  USTech Solutions, Noida, Uttar Pradesh   * Working on AI based Resume parsing, JD parsing, Resume matching, Resume Scoring Solutions using advance NLP techniques. * Deploying the model in AWS cloud using Docker Container, Rest API * Leading a team of 17 Data Scientist * Developing the POCs   Data Scientist (Machine Learning Architect)  HCL Technologies, Noida, Uttar Pradesh   * Successfully developed **abstractive Text Summarization** Model * Developed the customer’s **feedback Classifications** Model * Developed the **personalized recommendation Engine** and **cold start recommendation engine** * Developed a **Customer Segmentation** Model based on RFM (Recency - Frequency - Monetary Value) * Developed a classification model to predict the customers who can **upgrade mobile** * Developed a **Churn Prediction** Model * Leading a small team of Data Scientists * Responsible for doing POCs * Designing, coding and deployment of Data Science related Solutions * Demonstrated full expertise on Natural Language Processing, Machine Learning and Deep Learning |  |  |  |  | | --- | --- | --- | | **2016-06 - 2017-06** |  | Data Scientist  TATA Consultancy Services, Pune, Maharashtra   * Developed **Credit Risk model** by building Probability of Default (PD), Loss Given Default (LGD) and Exposure at Default (EAD) model to calculate the Expected Loss (EL) * Developed a classification model to predict the **recurrence of hepatitis C disease** after the completion of treatment. * POCs developments |  |  |  |  | | --- | --- | --- | | **2005-03 - 2016-06** |  | Senior Developer Analyst  Aon Hewitt (India) Pvt. Ltd, Noida, Uttar Pradesh   * Developed a classification model to predict the **loan repayment**. * Developed classification model to predict whether person can be **prospective Bank Customer**. * Developed a **sentiment analysis** model for comments written by employee in yearly engagement survey. * Developed an **association rule mining** model to find patterns in the variables - which accident conditions frequently occur together. * Developed a **regression model** to predict the emission of various cars as per Environmental Protection Agency (EPA). |  |  |  | | --- | --- | |  | **Certifications** |  |  |  |  | | --- | --- | --- | | **2019-11** |  | DAT203.1x: Data Science by Microsoft [Link to Certificate](https://courses.edx.org/certificates/46042e18cf14459ea0e9b583a57503f4) |  |  |  |  | | --- | --- | --- | | **2016-12** |  | Deep Learning 101 BY IBM [Link to Certificate](https://courses.bigdatauniversity.com/certificates/9f4eac86f09b4c19b452c1ef19be7d8a) |  |  |  |  | | --- | --- | --- | | **2017-01** |  | Data Science Workshop by IIT Bombay |  |  |  | | --- | --- | |  | **Publications - Research Papers** |   **Title: Imbalanced big data classification: A distributed implementation of SMOTE** Library: ACM DL Digital Library Conference: 19th International Conference of Distributed Computing and Networking Conference URL: https://iitbhu.ac.in/ICDCN2018/workshops.html Paper URL: <https://dl.acm.org/citation.cfm?id=3170535>   |  |  | | --- | --- | |  | **Blogs** | |  | |  |  | | --- | --- | |  | **Contact** |   Address  1002, T-12, Unitech Horizon  Greater Noida, Uttar Pradesh  Phone  +91 9891223932  E-mail  zamir.siddiqui@outlook.com  LinkedIn  <https://www.linkedin.com/in/zamir-ahmad-siddiqui-7>  Github:  <https://github.com/settings/repositories>   |  |  | | --- | --- | |  | **Skills** |   **Natural Language Processing**: NER, POS tagging, TF-IDF, Text Summarization, Text classification, Sentiment Analysis, word embedding- word2vec, glove, BERT, Flair, Allennlp, Spacy, Rasa    Excellent  **Machine Learning**: Linear Regression, Logistic Regression, Decision Tree, Random Forest, SVM, KNN, GBM, XGBoost, Naive Bayes apriori, K-Means, Customer Segmentation RFM, LVP, Expected loss (EL) - PD, LGD, EAD    Excellent  **Deep Learning**: ANN, CNN, RNN, LSTM, Seq2Seq, Attention,  Keras, Tensorflow, Pytorch    Very Good  **Tools**: H2O, R, Python, SparkML, google-colab, Flask, StreamLite, AWS EC2, Docker, Rest API    Excellent  **Domains:** e-com, Healthcare and Pharmaceuticals, Banking and Finance, Insurance, Telecom |
|  |  |  |

* <https://rwithzamir.wordpress.com/>
* <https://www.linkedin.com/posts/zamir-ahmad-siddiqui-776773110_streamlit-activity-6638476682053419008-igku>
* <https://www.linkedin.com/posts/zamir-ahmad-siddiqui-776773110_zamirsiddiquitext-classification-using-bert-activity-6639805564069220352-nrNp>

|  |  |
| --- | --- |
|  | **Live Application** |

<https://summary-ner-app.herokuapp.com/>