Supratim Haldar

Bangalore, India
Phone: +91 8884688864

Email: supratimh@gmail.com **Website:** https://supratimh.github.io
LinkedIn: https://www.linkedin.com/in/supratim-haldar-232bb16 **GitHub:** https://github.com/SupratimH

- Working as **Manager – Data Science** at **Head Digital Works Pvt. Ltd.**, a pioneer in online gaming with a rich portfolio of cricket.com, fanfight.com and ace2three.com.
- Have **6 years of hands-on experience in Data Science/Machine Learning** and 10 years in Software Application development and consulting.
- At present, leading the data science team of **cricket.com** - focused on Criclytics, and consumer and marketing analytics.
- Sound knowledge of **Statistics and Probability Theory** and **Machine Learning Algorithms**.
- Insightful experience in sports (especially cricket), telecommunications, online retail and any digital-service-provider business domains.
- Experienced in working as part of both **Product Development/Engineering** and **Project Implementation/Consulting** teams across organizations.
- Good at building large-scale applications intertwined with Predictive Analytics capabilities, catering to both consumer and enterprise markets.
- Experienced in **leadership roles** with managing teams of 10+ members, mentoring and people development.

PROFESSIONAL EXPERIENCE

HEAD DIGITAL WORKS PVT. LTD. **Bangalore, India**
Manager – Data Science **September 2019 - Present**

- At present, leading the Data Science team of cricket.
- Designed and developed the following key data-driven features:
  - **Criclytics** prediction engine (https://www.cricket.com/criclytics) – an unique feature among cricket related applications
    - Used by **35% of users** of the app.
    - Built a **ML classifier and MC simulation** model to predict **probability of win for teams**, with an accuracy of 75% on pre-game and 85% on in-game scenarios. The accuracy is among the best in industry.
    - Built a **regression and MC simulation model** to predict runs and wickets of players and teams before and during a game.
    - Built a **rating algorithm based on Elo Rating System** to calculate strength of teams and players, on which the above ML models are developed.
    - Built rule based engine to generate post-game insights like Ranking of Players and Game Changing Overs.
    - Built simulation based model to predict which teams will qualify for play-offs.
  - **Fantasy Research Center** (https://www.cricket.com/fantasy-research-center) – a feature to assist fans in forming teams for daily fantasy cricket (like Dream11, Fanfight etc.)
    - Results in **15% of app usage**.
    - With linear programming optimization method generate 6-7 teams programmatically, which ranks within top 30 percentile in fantasy leagues.
    - Fetch, calculate and present relevant data required for creating fantasy teams.
  - **Article Recommendation System** – a **content-based recommendation** framework

        - Developed from scratch by converting content of articles into TF-IDF vectors and calculating the distance between term vectors with Cosine Similarity.
        - Deployed as an API and calculates similar articles as and when a new article is published.
        - **Increased articles read on the platform by 12%**.
    - **Worked with engineering team in design and development of end-to-end pipeline including deployment of the models as Flask based APIs for all the above models.**
- User and Marketing Analytics
    - Designed ETL process to export data **from Clevertap to in-house MySQL DB**.
    - Design experiments and **perform A/B and Hypothesis Testing** to analyze behavior of app users.
    - **Increased DAU by ~11%** with targeted campaigns to segments of users, by segmenting users with RFM analysis on set of features.
    - Developed model to **predict churn on Day-1** with ~71% recall.
    - **Reduced Day-1 by ~15%** with recommendations to product and marketing teams based on feature analysis from churn prediction model.
    - Work very closely with Product, Tech/Engineering and Marketing teams to implement changes required for user retention and engagements.

ORACLE INDIA **NYC and St Louis, USA and Bangalore, India**
Solution Architect of apps and analytics products (IC Role) **April 2009 – September 2019**

- Worked as a Lead Architect of a **Hybrid Cloud Solution with Predictive Analytics Capability for Communication Industry** (CX on cloud, and order orchestration, service fulfillment, monetizing and analytical functionalities on premise).
    - Responsible for **designing end-to-end architecture**, **business process framework** and **design of M/L framework**.
    - Designed Machine Learning Model for following use cases
        - **Cluster Analysis** to generate segmented profile of customers
        - **Classification** model to predict whether a customer might **churn,** with accuracy of ~80% on real-world data

AMDOCS DEVELOPMENT LTD. **Limassol, Cyprus**
Senior Subject Matter Expert **July 2005 - February 2009**

- Worked as a Development Team Lead for both **Amdocs Engineering and Implementation Teams**

USHACOMM INDIA PVT. LTD. **Kolkata, India**
Software Developer **June 2003 – June 2005**

- As part of **Core Development Team,** was involved in design and development **Unicorn 6.0** product, mainly focused on **Inventory Management**, **Dealer Management** and **System Administrator** modules.

KEY SKILLS/KNOWLEDGE

- **Programming:**
    - **Proficient:** Python (including Numpy, Scipy, Sklearn, Pandas, Matplotlib, Seaborn)
    - **Intermediate:** R (including tidyverse, tidytext)
- **Database:** SQL, PL/SQL, NoSQL (MongoDB)

- **In-depth mathematical and hands-on understanding of M/L algorithms**:
  - Regression and Classification
  - CART and Decision Forests and other Ensemble Methods
  - K-Means Clustering, PCA, Anomaly Detection, Collaborative-Filtering and Content-Based Filtering Recommenders
- Data Ingestion (including web scraping), Data Visualization, Exploratory Data Analysis, Feature Engineering, Predictive Modelling
- **Basic NLP:** Data cleanup and preprocessing, POS tagging, Bag of Words vectorization
- **Methodology:** Waterfall, Agile and Scrum, Business Process Modelling
- **Tools:** MS Project, Confluence, GIT, SVN, CVS, JIRA, Quality Center etc.

PERSONAL PROJECT PORTFOLIO

A snapshot of few projects is provided below. Complete details of all projects are listed in http://supratimh.github.io

- A complete end-to-end project: Development of a classifier to predict result of ODI cricket matches, before the match begins. Prediction accuracy on matches of ICC World Cup 2019 outperformed Google's prediction accuracy (mine 79% vs Google's 75%).
  Also, predicted the top 4 teams accurately before even half of the matches were played.
- Cluster analysis of batsmen in CPL
- Developed datasets through web-scraping - AllSides : Ratings of bias in electronic media and BikeWise.org : Bike related reported incidents

EDUCATION

- University of Kalyani, West Bengal, India - B.Tech (IT), May 2003
- **Machine Learning** by Stanford University (through Coursera)
- **Neural Networks and Deep Learning** by deeplearning.ai (through Coursera)

EXTRA-CURRICULAR ACTIVITIES/ADDITIONAL SKILLS/HOBBIES

- Photography (https://500px.com/supxyz)
- Sports: Cycling, Running, Badminton, Table-Tennis
- Reading fiction, non-fiction, biographies and autobiographies