ASTHA PURI
(612) 481-5976 | asthashie@gmail.com | New York City, NY | [LinkedIn](#) | [YouTube](#)

SUMMARY

Seasoned Data Scientist specializing in responsible AI practices, driving impactful projects in healthcare, customer data platforms, and machine learning pipelines. With a track record of architecting innovative recommendation systems, implementing real-time search enhancements, and leading teams to optimize user experiences, I bring a unique blend of technical expertise and leadership. Let's collaborate to propagate responsible AI practices and empower future AI professionals.

SKILLS

SQL (Snowflake), Python (Scikit-learn, PyTorch, Tensor Flow), Neural Networks, R, PySpark, Statistics, DBT Labs, Machine Learning, Predictive Analytics, Natural Language Processing (NLP), Natural Language Understanding (NLU), Recommender Systems, Time Series Forecasting, Looker, Tableau, Segment, Amplitude, Mixpanel

EXPERIENCE

CVS HEALTH   New York, NY
Senior Data Scientist   Apr 2023 – Present
- Architected innovative 'You May Also Like' recommendation system using Word2Vec embeddings, fostering 30% higher click-through rates by deciphering user interactions into enriched feature vectors for precise content suggestions. Applied meticulous data preprocessing and hyperparameter tuning for optimal model performance and seamless platform integration.
- Innovated a successful Generative AI-driven POC for succinct product review summaries and sentiment analysis on the website. Leveraged advanced NLP models, resulting in 20% faster review comprehension and enhanced customer satisfaction through informed purchasing choices.
- Implemented real-time search autocomplete using RNNs on the company website, boosting search efficiency by 40% via data-driven modeling and learning from past search behaviors.
- Led and supervised junior data scientists in projects focused on enhancing the website's search experience. Coordinated project scope, milestones, and resources, resulting in the successful implementation of search optimization strategies that improved user engagement and satisfaction.

TWILIO SEGMENT   Los Angeles, CA
Senior Data Scientist   Feb 2022 – Feb 2023
Led ML and AI for Segment's Customer Data Platform(CDP), driving growth, profitability, and platform health. Collaborated with product managers, researchers, engineers and analytics leaders to deploy scalable ML solutions across the organization
- Saved $240k annually by forecasting traffic for migrating customer's data to EU following GDPR and data privacy regulations, optimizing compute and storage resource allocation.
- Improved operational efficiency and reduced latency by 30% by designing and building analytics pipelines in DBT
- Grew demo rate by 7% by designing and running experiments across the customer acquisition funnel, to evaluate effectiveness of different marketing strategies and support research
- Developed a resource/cost-profit framework for finance and leadership to understand cogs and impact of initiatives on the bottomline forecasted to help them prioritize and build roadmaps on yearly initiatives (expansion, churn, operational/support cost, retention)
- Designed and launched success and operational metrics of products ([HIPAA](#) and [Profiles Sync](#)) across their pilot, beta and GA stages to understand customer adoption and make go/no-go decisions

ORACLE   Broomfield, CO
Lead Data Scientist - Machine Learning   Aug 2021 – Feb 2022
ML and AI owner for multiple product development work streams - KPI definition and data instrumentation design, ad-hoc and executive reporting and building and deploying Machine Learning pipelines
- Improved customer retention by 9% by predicting and flagging customer SaaS subscriptions for Fusion Applications, that were likely to churn in next 9+ months using classification models, enabling support and sales team to take actionable steps to improve customer retention
- Reduced escalation handling costs by 15% by deploying MIT's DeepMoji for customer sentiment analysis, predicting service request escalation using ensemble models in PySpark

Data Scientist - Machine Learning   July 2018 – July 2021
- Built data pipelines that processed stream events using [Kafka compatible services](#) (with OSaK and Spark applications) to provide real-time analysis of streaming data
- Analyzed time series data from server logs, studied trends across KPIs (restarts, downtimes, SLOs) and built web applications and visualizations for leadership, securing application authentications and data.
- Performed statistical testing (t-test, chi sq test, etc.) to understand product performance and identify areas of improvement.

CARLSON ANALYTICS LAB   Minneapolis, MN
Machine Learning Consultant   June 2017 – May 2018

- Forecasted daily footfall in Nickelodeon Universe amusement park for 2018 using time series forecasting, decreasing error rate in capacity planning by 12%
- Optimized promotion strategy to increase market share by building region-level fractional logit models and developed randomized experiments (A/B tests) to measure campaign impact.

FORD MOTOR COMPANY   Chennai, India
Business Intelligence Engineer   July 2015 - April 2017
- Built self-service BI reports in Tableau, QlikView and SAP BO for global car parts, maintenance, warranty, and pricing teams. Performed deep dive analysis and presented business insights to executive management.

EDUCATION

UNIVERSITY OF MINNESOTA, Carlson School of Management, MN
Master of Science in Business Analytics (MSBA)   *May 2018*
Coursework: Prescriptive and Customer Analytics, Machine Learning, Time Series and Heuristics

VIT UNIVERSITY, India
Bachelor of Technology - Electronics and Communication (ECE)   *May 2015*

LEADERSHIP

- Book reviews completed:
    - Reliable LLMs by Rusheel Shahani (Manning Publication)
    - Hugging Face in Action by Wei-Meng Lee (Manning Publication)
    - Generative AI for Web Development (Springer Nature Publication)
- Book chapters published:
    - Puri, A., Nayar, P., Bamel, P., Sindhu, B., Puri, A. (2024). Implementing Distress Screening and Psychological Assessment at the A.R.T. Clinic. Pub. by Indian Fertility Society (Ed.), Counselling in Assisted Reproduction Technology (A.R.T) (pp. 63).
    - Bamel, P., Puri, A., Nayar, K. D., Singh, T., Jogy, S. (2024). Use of Interest-Based Technology for Patient Care. Pub. by Indian Fertility Society (Ed.), Counselling in Assisted Reproduction Technology (A.R.T) (pp. 83).
    - Puri,A, Nayar,P., Bamel, P.,Puri,A and Kumar,A (2024) Invisible Grief During A.R.T.. Published by Indian Fertility Society (Ed). Counseling in Assisted Reproduction Technology (A.R.T.) pp 23
- Papers published:
    - [Internet of Things (IoT) based Attendance and Intrusion Detection System](#)
    - Bamel, P., Sindhu, B., Sindhu, S., Puri, A., Singh, T. (2024). Recent Eclectic Approach to Psychotherapeutic Interventions in the Indian Context - Subconscious Energy Healing Therapy (S.E.H.T). *International Journal of Creative Research Thoughts (IJCRT)*, 12(2), a434-a442. ([Link](#))
    - Puri, A., Bamel, P., Sindhu, B., et al. (2023). Recent advances in psychotherapy in the Indian scenario - subconscious energy healing therapy: S.E.H.T for infertility counselling. J Psychol Clin Psychiatry, 14(6), 182–194. DOI: 10.15406/jpcpy.2023.14.00750 3. ([Link](#))
    - Puri A, Sindhu BD, Puri A, et al. Hypnotherapy as an intervention in infertility treatment. Art Human Open Acc J. 2023;5(3):214–218. DOI: 10.15406/ahoaj.2023.05.00211 ([Link](#))
    - Banerjee P, Sindhu BD, Sindhu S, et al. Exploring the intersections of AI (Artificial Intelligence) in psychology and astrology: a conceptual inquiry for human well-being. J Psychol Clin Psychiatry. 2024;15(1):75–77. DOI: 10.15406/jpcpy.2024.15.00761 ([Link](#))
    - Puri A, Banerjee P, Nayar P, et al. Understanding the impact of environmental pollutants on infertility counselling: insights from the Indian scenario. Art Human Open Acc J. 2024;6(1):35–37. DOI: 10.15406/ahoaj.2024.06.00218 ([Link](#))
    - Puri,A, Navya,N and Shammi (2019). Malaise of Domestic Violence: Scarring Children's Well Being ([Link](#))
- Featured on [spotlight](#) for Enterprise Generative AI Summit West Cost ([Resource Link](#))
- Patent on LLM evaluation in filing process
- Professional Membership - [Indian Fertility Society](#)
- Speaker at Women Who Code events:
    - [Choosing the right metric for your business](#)
    - [SQL Intermediate 1](#)
    - [SQL Intermediate 2](#)
- Speaker at DSS LinkedIn - [Increasing customer satisfaction in B2B using ML](#)
- Speaker at FertiVision 2023, Gurgaon, India
- [Technical writer](#) at Towards AI
- Mentor at [Women in Analytics](#)
- Volunteer:
    - University of Minnesota and University of Southern California students with interview prep

- Animal fostering
- Bhumi India (NGO) - taught robotics to students