Indrajit Singh

Lead Data Scientist

Enthusiastic Software Engineer having 10+ years of experience. Eager to contribute to team success through smartwork, attention and excellent organizational skills. A clear understanding of NLP and Deep Learning concepts. Data Engineer veteran. Motivated to learn, grow and excel in Technology.

Contact : +91 9591182659   **email : indrajitsingh37@gmail.com**   **Address: Pune, MH-411057**

in ⬡ Ⓗ M**Portfolio**

Work History [Sep-2012 – present]

2021-04 - Current

Lead Data Scientist

eClerx, Pune, Maharashtra

- Leading VoiceAi and TextAi solutions at eClerx as an SME for the past 1 year.
- Brought current speech to text model accuracy from **67% to 89%** using better data quality audit and model training.
- Optimized AutoQa code performance from 1 chat file in 1.5 **minutes to 5 seconds**.
- Built a POC and won **1M deal** on ecommerce recommendation engine for product SKUs on client's top competitors.
- Managing & guiding a **team of 6+** Data Scientists/Analysts on critical problem solving on AutoQa, Speech and general model building approach for VoiceAi.
- Providing consultations for more than **30+ projects** at eClerx.

2018-12 - 03-2021

NLP Research (Unpaid-Professional Learning)

CellStrat, Bangalore, KA

- Collaborated with Cellstrat AI Lab on AI and NLP Research methods as Intern for 2+ Years.
- Built state of art models using Deep Learning such as Questions Answering System, Text Summarizer, and TTS.
- Developed and implemented high-performance QA models having more than 93% using ALBERT and BERT based models.
- Recently published research paper on BERT - Algorithm Behind Better Search Engines under CellStrat.
- Built Abstract Text Summarization using LSTM. Worked on NER on CRM log data using CRFsuite and LSTM.
- Worked on a Kaggle notebook for GitHub Bugs Prediction using BERT.
- Help mentoring several Data Scientists at CellStrat and built connections.

Top Skills

- Compassion ★★★
- NLP ★★★
- Data Science ★★★
- PySpark ★★★
- Data Engineering ★★★
- SQL ★★★
- Python ★★★

Achievements & Awards

- NLP Hackathon champion at Cellstrat.
- Blogathon Winner @eClerx.
- Multiple Kaggle competition winner.

Certifications

- CellStrat Certified NLP expert.
- Accredited Informatica ETL Administrator and Developer Certification.
- RPA certification from University of Toronto.

2020-07 –
2021-04

Senior Customer Success Engineer

Fivetran, Bangalore, KA

- Established a Quality-driven approach to deliver excellent customer satisfaction by ensuring three key principles are in place: Understand | Empathize | Solve
- Achieved 100% customer satisfaction score in consecutive quarters.
- Able to achieve < 90% SLA for resolving customer issues on connectors.
- Built an excellent understanding of data pipelines on cloud, dbt, data flow, and model data warehousing principles.

2018-03 -
2020-06

Senior Software Engineer

Informatica Corp, Bangalore, KA

- Contributed to an outstanding and dynamic team of ETL Engineers and product specialists & significantly boosted the overall success of the organization.
- Experience in troubleshooting IDQ platform and Complex product issues deployed on Cloud servers.
- Used sentiment analysis for understanding customer communications and alerting end-users.
- Data Quality: Implemented ML Algorithms (Bigram, Edit Distance etc.), token labelling, and regular expressions to remove duplicates, nulls, and errors in data.
- Designed Data Quality course modules for new engineers in the team.
- Designed and implemented a complete ETL process using Informatica Power Center, Data Quality and BDM.

2012-09 –

2018-03

Application Engineer

Informatica Corporation, Bangalore, KA

- Learned the right approach to understand, dissect and debug any critical customer problem.
- Got to understand and reproduce product behaviors to determine problem root-cause(s), issue work-around and solutions.
- Worked on large enterprise applications on various operating systems (Linux/Unix, Windows etc.), databases (MySQL, MS-SQL, Oracle, MongoDB, AWS Redshift etc.) and ecosystems.
- Worked on NoSQL/SQL large data stores upon multi-stage data pipelines for resolving customer issues.
- Closely worked with customers to plan their server capacity including ongoing monitoring and server resource consumption to avoid any critical failure.

Top 3 Projects

- **VoiceAi** – Built an application from scratch for Speech-to-text and text analytics on the generated text. The solution is provided to several enterprise customers.
- **AutoQa**- A sentence similarity based, sentence transformer model which works on both keyword and meaning based phrase matching techniques.
- **Project Self-Realization**: I am working on one of my personal projects which will help us find answers such as Who I am, why we are born in this earth etc.

Education

2016-06 -
2018-08

M.Tech: Data Analytics
Birla Institute of Technology - Pilani

- Dissertation: Recurrent Neural Tensor Network for Predictive Escalation

2008-07 -
2012-07

B. Tech: Information Technology
West Bengal University Of Technology - Durgapur