

# SIDDHANT SADANGI

Data Scientist @Thomson Reuters | Technical Content Creator  
@Neptune.ai | Ex – Consultant @Deloitte

[siddhant.sadangi@gmail.com](mailto:siddhant.sadangi@gmail.com)

+91-789-946-8877

Bengaluru, India

[linkedin.com/in/siddhantsadangi](https://linkedin.com/in/siddhantsadangi)

[github.com/SiddhantSadangi](https://github.com/SiddhantSadangi)

## PROFESSIONAL SUMMARY:

Working as an ML Developer Advocate with **Neptune.ai**, one of the leading SaaS MLOps platforms, where I have been instrumental in **designing and drafting their documentation** and product comparison pages. Previously a Data Scientist and Full-Stack Developer with **Reuters** where I contributed to over **\$3M in Annual Recurring Revenue** through models, pipelines, and analyses on client and product engagement for the Editorial, Marketing, Product, and Sales teams. Also have ~4 years of ETL development experience with Deloitte.

## WORK EXPERIENCE:

### ML Developer Advocate

[Neptune.ai](https://neptune.ai), Remote | July 2022 – Present (full-time) | October 2020 – June 2022 (part-time)

- Maintaining and upgrading the [Neptune examples](#) open-source GitHub repository
- Optimised existing GitHub Actions to test all code on different OS and with different Python versions. **Increased code coverage to 100% while reducing the testing time from eight hours per workflow to less than one**
- Creating integration and how-to code [examples](#) and [documentation](#)
- Research and drafting for Neptune's [competitor comparison](#) pages.

### Data Scientist, Reuters

[Thomson Reuters](#), Bengaluru | June 2019 – June 2022

- Developed an ETL pipeline for detecting Image copyright infringements using third-party detections of Reuters' images on websites, and matching detection details to internal client databases to filter out false positives. The project delivers around **\$2.4M ARR**
- Enhanced an existing recommender system by adding a filtering model for increasing the relevancy of recommendations to Account Managers based on feedback received for previous recommendations. This process **decreased invalid recommendations from 50% to 10%**. Also implemented an email notification script using Amazon SES to each account manager whenever recommendations were generated for their accounts, with periodic follow-up reminders. This has brought **TAT down from 180 days to 20 days**
- Delivered analytical reports to sales and account managers related to their clients' usage trends which have contributed to **\$500K+ ARR**
- Lead the development of pipelines to **predict upselling opportunities to clients**. Delivered a **10% increase in revenue** from highlighted clients in 2020H2
- Achieved an **accuracy of 98% and 96% for English and Spanish** respectively for a text classification model on news stories
- Worked as a **full-stack developer** to enhance and maintain the Agency Customer Engine (ACE), a one-stop platform for the account, sales, and product managers to track client engagement and product performance
- Built a pipeline for **Anomaly Detection** on Reuters.com that utilizes Scikit-Learn's Isolation Forest Algorithm to assign an anomaly score to data points, and the points can then be flagged as an anomaly based on a dynamic threshold and granularity by the end-user through a Tableau dashboard
- Developed an NLTK PoC to **quantify gender diversity in news articles**

## TECHNICAL SKILLS:

### Development

#### Languages:

Python, SQL, JavaScript, Java, PHP, Shell, Visual Basic, Batch

#### Machine Learning:

Classification, regression, clustering, NLP, computer vision, anomaly detection, time-series forecasting

#### Visualization Tools:

Tableau, MS Excel, AWS Quicksight, Metabase, Matplotlib, Plotly, Streamlit

#### Databases:

Teradata, MySQL, PHP MyAdmin, Hive, AWS Athena, Apache Druid, Google BigQuery

#### ETL:

Informatica PowerCenter, Teradata, Unix, PuTTY

#### Version Control:

Git, GitHub, ClearCase, Neptune

#### IDEs:

- Hosted live hands-on **workshops on Text and Image Processing and Classification** for beginners in Machine Learning
- Founding member of the **Bangalore ML Community of Practitioners** , and thread-lead for the **Thomson Reuters Bengaluru Tech Council**.

## ETL Developer Consultant, Strategy & Analytics

[Deloitte](#), Bengaluru | August 2015 – May 2019

- Achieved **94% accuracy** on a **deep learning computer vision model using Python** which identifies plant diseases just from a photo of the leaves
- **Increased operational efficiency by 50% and reduction of FTE for weekend support by 100%** by optimizing processes, reducing process failures, and documenting standard operating procedures within the team
- Independently created Excel and VBA tools which were used across the firm and led to a **saving of upto 10% in Informatica PowerCenter development efforts**
- Managed manual and automation testing independently for extracts, with **0% spillover to UAT**
- Worked as a **Python developer** for a Text Extraction and NER PoC. Was responsible for creating a pipeline to extract and clean text from noisy scanned images
- Won an **Outstanding Award** for my work with the **Deloitte Machine Learning Guild's** Engagement and Expansion team for organizing monthly Machine Learning quizzes and interaction sessions within the firm. Also managed the drafting and publishing of the monthly guild newsletter.
- Won an **Applause** and a **Spot award** for ETL development and support for client projects
- Won a **Spot award** for contributions and innovations towards the monthly service line newsletter.

## OTHER TECHNICAL CONTRIBUTIONS:

- Public apps:
  - [Used Car Picker](#) - ML-powered app that helps users pick a car with the highest resale profit potential based on their usage and preferences
  - [Wordle Solver](#) - has a **success rate of 96%** and takes an average of 4.2 attempts to converge to a solution for Wordle (hard mode, without knowledge of the 2315 possible wordle solutions)
  - [Image WorkDesk](#) – lets users crop images and play around with image properties like brightness, saturation, contrast, and sharpness. They can also randomize these properties and download the final image at the click of a button
  - [SpellBee](#) – Python-powered interactive command-line spelling bee app.
- Public **Tableau dashboards**
- Articles on [Medium.com](#) (published by **Analytics Vidhya** and **Better Programming**)
- Hosted a **live-coding workshop** on [Image Processing using Python](#) as a part of Analytics Vidhya's The DataHour series.

## VOLUNTEER EXPERIENCE:

- **Lead** the Thomson Reuters India Green Team. Won an **Ignite Award** for my contributions (2021-)
- Participated in Oxfam Trailwalker 2020, where I **walked 100 km in 37 hours** to raise funds for charity (February 2020)
- Employability skills mentor for less privileged students with Youth for Seva (2020)
- Volunteered as a journalist for a UNESCO-affiliated online magazine (2011-2014).

Jupyter Notebook/Lab, Google Colab, PyCharm, PhpStorm, IntelliJ, Eclipse, AWS Sagemaker Studio, VS Code

## Scheduling Tools:

Airflow, Oozie, Crontab, Control-M, WLM

## Environment Management:

Docker, Conda

## CI/CD:

GitHub workflows

## Collaboration Tools:

MS Teams, Slack, Jira, Sococo, ClickUp, Neptune, Google Workspace, Loom, MS Office Suite

## Technical

## Documentation:

GitBook, Markdown, Restructured Text, Sphinx

## SPOKEN

## LANGUAGES:

English – Professional Fluency  
Hindi – Native Fluency  
Odia – Proficient socially  
French – Beginner

## HOBBIES AND INTERESTS:

Environmentalism; listening to financial and data science podcasts; reading and writing technical articles on machine learning,

## EDUCATION:

### Bachelor of Technology, Electronics and Electrical Engineering

[\*Kalinga Institute of Industrial Technology\*](#), Bhubaneswar | 2011-2015 | 8.13 CGPA

- Student coordinator of Qutopia, the Quiz society of KIIT University
- Student editor of Kritika, the Annual Students' Magazine of KIIT University
- Chief of the International Press and member of the Organizing Committee of KIIT International Model United Nations Conference (MUN), one of the largest MUN conferences in the country
- Actively participated in, adjudicated, and drafted newsletters for more than 20 MUN conferences across the country
- Awarded with the '**Best Student Coordinator**' award for organizing Udghosh – the largest open quiz in the country.

python and analytics;  
building Streamlit apps;  
quizzing.