



# Nikhil Vidhani

*Lead Data Scientist, Target Corporation*

## Selected Past Projects

- 2023 – **Replay Mode and Logging, Target.**
- 2024
- Engineered a post-mortem analysis tool, employing massive parallelization of sequential tasks (*Replay*), for quick output evaluation
  - Implemented granular logging for granular scrutiny of allocation decisions. Used tabular representation for easy consumption of logs
  - Improved labor allocation accuracy from 50% to 90% and reduced Air service utilization by 10x
- 2022 **QA Analytics, WNS.**
- Highly configurable workflow tool designed to catch human errors in invoice indexing
  - Provides a one-stop solution for outlier and anomalous transaction detection
  - Automated bundling with portable-R, necessary packages, and configs for scalable hassle-free deployment
  - Single tool deployed across 5 different audit teams within WNS. Caught \$5+ Mn worth of scanned invoices with sanity issues.
- 2022 **Audit Mate, WNS.**
- ML solution to classify manual invoice-indexing errors and help controllers predict them right from the comfort of excel
  - Built and served using *h2o* framework and deployed as an API service
  - Increased audit error incidence from 10% to 95%
  - <https://github.com/nik141088/audit-mate>
- 2022 – **Invoice Processing, Personal.**
- Present
- Deep Learning based tool to compare similar looking invoices
  - Multi-tiered comparison based on image embeddings and OCR text embeddings
  - Can be used to boost data processor's efficiency and productivity
  - <https://github.com/nik141088/invoice-processing>
- 2020 – **Factiva Download, IIMB.**
- 2021
- Developed a robust, self-organizing, and highly parallel web-scraping system for Dow Jones Factiva news articles
  - Automated network/scraping/parsing failures, de-duplication, and storage management
  - Extracted and processed 25+ Mn articles and analyzed sentiment using GPT2/DistilBERT
  - <https://github.com/nik141088/factiva-download>