

Nischal Mahaveer Chand

Boston, MA | nischal.bhatewara@gmail.com | (857) 277-9145
Blog: medium.com/@darkestfloyd | LinkedIn: nischalmc | GitHub: DarkestFloyd

WORK EXPERIENCE

BERG LLC, Framingham, MA

Nov 2019 - present

Data Scientist II

Tools: R, RStudio, Shiny, tidyverse, git, linux, ganglia

- Statistical analysis (regression analysis, survival analysis, hypothesis testing, FDR correction):
 - Perform survival analysis on cohort of $n > 100$ pancreatic cancer patients to understand effectiveness of various chemotherapy and adjuvant treatments using Kaplan-Meier curve in R, resulting in a poster presentation at ESMO 2020.
 - Perform statistical analysis on omic data ($p > 20k$) using Wilcox tests and ordinal regressions to obtain predictive biomarker panel for early detection of Pancreatic Cancer and for survival related outcomes in cohort of $n > 100$ patients.
 - Perform validation study using open and freely available data from NCBI websites, backing internal results with independent data resulting in stronger papers and posters for conferences.
- Data pipelines (ingest RESTful API (JSON), SQL databases, excel sheets, webscraping; orchestration using CRON):
 - Built automated data processing pipeline to process longitudinal clinical data of $n > 200k$ patients, resulting in discovery of various clinical features that impacted survival outcomes for COVID-19 patients by bAlcis®.
 - Built data processing pipeline to gather data from multiple data sources - clinical data using RESTful API, omics data from SQL servers and excel files - for $n > 450$ patients, further performing QC checks and merging data based on timepoints.
 - Built pipelines to obtain external Patent data (uspto.gov), DrugBank data (drugbank.com), PubMed data (NCBI) and cross-reference with internal biomarkers for use by upper management, cutting down time required to find key markers.
 - Improve legacy SQL data pipeline by applying filters and aggregation in SQL to cut down processing time by 68%.
- Data dashboards and visualizations (R Shiny dashboards; ggplot2 visualizations):
 - Built interactive data dashboard to allow clinicians to examine impact of clinical features on survival outcomes for various subpopulation of COVID-19 patients (by age, race, comorbidities, etc), further deployed dashboard on Azure cloud.
 - Built R-Shiny dashboards to track key metrics and provide overview of internal processes and data stores.
 - Create data visualizations of analysis methods and results for publications and posters using ggplot2 in R.
- Deep learning: Trained over 2k deep neural network models to optimize running time and precision of bAlcis® using keras, cutting down on-boarding time of new team members getting started with bAlcis®.
- Provide support to Analytics team as Linux administrator for Analytics Server (Turing) and HPC cluster of 55 CPU nodes and 1 GPU node (Bayes cluster), further deploy and test new versions of bAlcis® on cluster.
- Contribute to company IP via software implementations and improvements to AI platform and legacy dashboards.

Marcus Institute for Aging Research, Hebrew SeniorLife, Roslindale, MA

July 2018 - Dec 2018

Junior Data Scientist / Co-op Student

- *AD Supplement*:
 - Performed regression analysis and hypothesis testing on clinical trials data using R and SQL.
 - Actively communicated results to directors and researchers.
- *Smartphone Lab*:
 - Built data processing pipelines for time-series sensor data using R and SQL.
 - Improved data processing speeds by upto 13%.
- *shinyMRI*:
 - Built scalable R shiny application to visualize 3D and 4D MRI images.
 - Improved diagnosis speed for in-house Alzheimer's research.
 - Application received honorable mention by RStudio, Inc.

EDUCATION

NORTHEASTERN UNIVERSITY, Boston, MA

Khoury College of Computer Sciences, GPA: 3.79/4.0

Sept 2017 - Aug 2019

Master of Science in Data Science

TA positions: Data Management in R (Spring 2019); Unsupervised Machine Learning (Summer 2019)

ALLIANCE UNIVERSITY, Bengaluru, India

2013 - 2017

Bachelor of Technology in Computer Science and Engineering

PUBLICATIONS

- Kiebish, M.A., Tekumalla, P., Ravipaty, S. et al. Clinical utility of a serum biomarker panel in distinguishing prostate cancer from benign prostate hyperplasia. Sci Rep 11, 15052 (2021). <https://doi.org/10.1038/s41598-021-94438-4>
- Miller, Gregory M., et al. "Ondansetron use is associated with lower COVID-19 mortality in a Real-World Data network-based analysis." medRxiv (2021).
- Moser, A. J., et al. "1485P Identification of novel protein biomarkers for FOLFIRINOX-based chemotherapy response in advanced pancreatic adenocarcinoma using patient omics and Bayesian AI." Annals of Oncology 32 (2021): S1094.
- Moser, A. J., et al. "1547P Project Survival®: High-fidelity longitudinal phenotypic and multi-omic characterization of pancreatic ductal adenocarcinoma (PDAC) for biomarker discovery." Annals of Oncology 31 (2020): S947-S948.