



SWETA RAI

Lethbridge Canada
sweta.raai@uleth.ca / 3683995619

SUMMARY

I'm a dynamic PhD student at the University of Lethbridge, with hands-on experience in data-driven biological research and machine learning. I enjoy exploring ways to improve model performance through thoughtful feature selection and hyperparameter tuning. With a curious mind and a strong research drive, I see learning as a continuous journey that keeps me inspired to uncover meaningful insights from complex datasets.

SKILLS

- Web scraping and data extraction
- Machine learning and scikit-learn
- Data manipulation with pandas
- Python and R programming
- Transcriptomics and gene expression analysis
- Methyloomics data analysis
- Web application development with Shiny
- Language fluency
- Complex problem solving
- Punctuality and work ethic
- Presentation development
- Document scanning
- Task initiative
- Computer proficiency
- Research skills
- Positive attitude
- Time management

EXPERIENCE

Mentoring Experience / University of Lethbridge, Biological science - Lethbridge, Alberta 01/2025 - Current

- Guided three students through academic and personal development.
- Monitored and evaluated mentees' progress, providing feedback and guidance to ensure success.
- Engaged consistently with mentees through designated contact points using various methods.
- Encouraged open dialogue between mentor and mentee pairs to promote trust, respect and understanding of different perspectives.
- Documented mentee progress and provided insights for future planning.

Phd Student / University of Lethbridge - Lethbridge, Canada 01/2024 - Current

- Conducted multi-omics data analysis for biomarker discovery using various integration methods.
- Performed differential gene expression analysis utilizing the "DESeq," "Limma" packages in R.
- Experienced in designing and applying hypothesis testing, correlation, regression, and classification analyses to uncover relationships and patterns in complex biological datasets.
- Demonstrated expertise in descriptive analysis by computing measures of central tendency and dispersion to characterize multi-omics datasets, revealing skewness and outliers in methylomics and transcriptomics profiles for rheumatoid arthritis biomarker studies.
- Designed R-based boxplots, violin plots, and PCA scatterplots for NGS data exploration, uncovering variability and outliers that optimised preprocessing pipelines for machine learning models.
- Applied goodness-of-fit tests (chi-square, Kolmogorov-Smirnov) to validate distributional assumptions in datasets, ensuring robust preprocessing.

- Demonstrated proficiency in contingency table analysis by interpreting chi-square statistics to identify associations between categorical variables.
- Performed two-sample t-tests and ANOVA for multi-sample comparisons to evaluate differences across experimental groups in bioinformatics datasets.
- Computed Pearson and Spearman correlations to quantify relationships between methylation patterns and gene expression levels across multi-omics datasets.
- Demonstrated an understanding of linear regression by modeling relationships between predictor variables and continuous outcomes using least-squares estimation to best-fit lines, and assessing assumptions like linearity, normality, and homoscedasticity.
- Implemented both supervised and unsupervised machine learning algorithms, including logistic regression, SVM, XGBoost, random forest, and deep neural networks.
- Applied feature selection techniques such as 'Boruta' and 'LASSO' to enhance model performance.
- Executed data pre-processing techniques, including normalisation, missing value handling, outlier detection, and hyperparameter tuning.
- Working on a user-friendly web application for analysing multi-omics data related to rheumatoid arthritis.
- Scraped data from NCBI GEO datasets and extracted PubMed IDs using BeautifulSoup and Pandas.
- Generated synthetic biomedical images using Generative Adversarial Networks (GANs) based on Kaggle datasets.
- Completed day-to-day duties accurately and efficiently.

Data Science Engineer / SirpiData Science - Bangalore, India

03/2024 - 10/2024

- Contributed to project at IISc Bangalore focused on machine learning model development.
- Built and tested multiple machine learning models for accuracy and performance.
- Fine-tuned models using hyperparameter tuning techniques, including MinMax and Standard Scalers.
- Researched problem statements to inform project direction and solutions.
- Presented data stories to clients through a web application, enhancing communication.
- Collaborated with team members to devise effective engineering solutions.
- Conducted research on materials and technologies to drive project enhancements.
- Maintained engineering records and supported project documentation consistently.

Data Science Intern / DataMites - Bangalore

08/2022 - 01/2023

- Built machine learning models for classification and regression tasks, achieving high accuracy metrics.
- Conducted exploratory data analysis on large datasets to uncover trends and patterns.
- Collaborated with team members on data cleaning and preprocessing to enhance dataset quality.
- Created visualizations using Tableau, Power BI, and R Shiny to effectively communicate findings.
- Documented methodologies and workflows in Jupyter Notebooks to ensure reproducibility.
- Evaluated model performance using metrics such as accuracy, precision, recall, and F1 score.
- Assisted in developing predictive models that analyzed customer behavior through various techniques.
- Engaged in peer reviews of analyses to maintain rigorous quality standards.

Trainee Medical Reviewer / Omega Healthcare Management Pvt Ltd - Bangalore

01/2021 - 10/2021

- Reviewed nearly 300 Non-small cell lung cancer patients' charts to ensure accurate data collection.
- Managed multiple tasks in a computerised environment to enhance operational efficiency.
- Conducted research on medications while maintaining HIPAA, PHI, and confidentiality standards.
- Focused on achieving production and quality deadlines consistently.
- Assisted in daily operations, providing support to senior team members as needed.
- Participated in team meetings to gain insights into operational processes and improvements.

- Learned and adhered to safety protocols during training sessions to ensure compliance.

EDUCATION

PhD: Bioinformatics

01/2024

University of Lethbridge - Lethbridge, AB, Canada

- Participated in the University's Mentorship Program.
- Supervised by Dr Igor Kovalchuk.
- Focused on multi-omics, data science, bioinformatics and web development

Master of Science: Applied Genetics And Bioinformatics

04/2022

Bangalore University - Bangalore

- Academic Achievements and Honours: Gold Medal
- Minor in: Biotechnology
- 8.7/10 GPA
- Thesis or Dissertation
- Title: Investigation of anti-arthritis compounds from *Cinnamomum zeylanicum* through in-silico methods.

ACHIEVEMENTS

- Received the Gold Medal award and 1st rank award by the Vice Chancellor of Bangalore University in the Master's degree (2022).
- University third rank holder in Bachelor's degree (2019).
- Certified Data Scientist (2023).

PERSONAL INFORMATION

Date of Birth: 12/18/98

LANGUAGES

English:

Professional

ACTIVITIES

Hiking