Umanga Bista

📍 Canberra, Australia   📞 (+61) 416 119 223   ✉ umanga.bista@outlook.com

🏠 bistaumanga.com.np | 🔗 /bistaumanga | ⌀ /bistaumanga

Research Interests

Applying machine learning and deep learning to solve real world problems in text mining, anomaly & fraud detection, predictive analytics, web & social data analysis, etc.

Education

PhD Candidate in Machine Learning Australian National University, Canberra   [03/2017 – 01/2021]
○ *Comparative Summarization of Document Collections*: We formulate the problem of comparative summarization as competing binary classifiers. We develop unsupervised and supervised models , extrinsic evaluations within this formulation. We show the applications in existing, and controversial news dataset that we curated.
○ Advisors: Prof. Lexing Xie, Aditya K. Menon, Alexander P. Mathews

Bachelors in Computer Engineering, Institute of Engineering, Pulchowk, Nepal   [11/2009 – 09/2013]

Experience

Research Assistant Australian National University, Canberra   [09/2018 – now]
○ Developing and maintaining 24×7 fault-tolerant and scalable data collection (tweets, news articles and images), processing and storage systems. Part of this data is used in my thesis.
○ Technologies: Python, Kafka, Redis, Elasticsearch, Logstash, Kibana, DBPedia Spotlight, Shell-script

Tutor Australian National University , Canberra   [07/2017 – now]
○ Designing lab modules, conducting tutorials, grading assignments and exams
○ Courses: Programming for Data Scientists[2020], Intro to Machine Learning[2020], Document Analysis[2017,18]

Research Engineer LogPoint, Lalitpur, Nepal   [10/2013 – 02/2017]
○ Developing algorithms, implementing, testing and maintaining for security analytics in LogPoint product
○ Projects: anomaly detection, streaming log clustering, streaming data analytics, alerts frameworks, etc
○ Installing and managing clusters of spark, kafka and secure HDFS, and helping interns in their projects
○ Technologies: Python, R, Java, Scala, Kafka, HDFS, Spark

Skills

Programming	Python, Scala, Java, Shell-scripting | familiar: MATLAB, HTML/CSS, C/C++
Data Science	R, Spark, PyTorch, Scikit-Learn, SQL, Regex, Elasticsearch stack, Hadoop, Kafka, Redis

Publications

[1] SupMMD: A Sentence Importance Model for Extractive Summarization using Maximum Mean Discrepancy, *U Bista, A Mathews, A Menon, L Xie* [ACL Findings: EMNLP '20]
[2] Comparative Document Summarization via Classification, *U Bista, M Shin, A Mathews, A Menon, L Xie* (AAAI'19)
[3] Comparative Summarisation of Rich Media Collections, *U Bista* [WSDM '19]
[4] Visualizing Graph Differences from Social Media Streams, *M Shin, D Kim, JH Lee, U Bista, L Xie* [WSDM '19]
[5] Knowledge Guided Attention and Inference for Describing Images Containing Unseen Objects, *A Mogadala, U Bista, L Xie, A Rettinger* [ESWC '18]
[6] Intelligent Clustering Scheme for Log Data Streams, *B. Joshi, U. Bista, M. Ghimire* [CICLING '14]

Referees

Available upon request.