🏠 300 W Franklin St,
Apt 301W,
Richmond VA 23220

DARSHINI MAHENDRAN

📱 (804) 349-6327
✉ samanthamahendran6@gmail.com
⌨ SamMahen
in/sammahen

Research Interests

My research interests lie in the area Natural Language Processing (NLP) and Data science. My main research focus is in Information extraction including Relation Extraction (RE), Named Entity Recognition (NER). I'm also interested in the Machine learning and Deep learning based research.

🎓 Education

Richmond, VA, USA   **Virginia Commonwealth Univ.**   📅 **May 2022**

- Pursuing Ph.D. in Computer Science (concentrated in NLP).
- **Advisor**: Dr. Bridget McInnes
- **Most Relevant Coursework:** NLP, Advanced Algorithms; Machine Learning, Fuzzy Logic Algorithm; Parallel Algorithms; Data mining and knowledge Discovery, High Performance Systems, Regularization methods for machine learning, Software analysis and testing.

Colombo, Sri Lanka   **Univ. of Peradeniya**   📅 **June 2015**

- B.Sc. (Hons) Computer Science (special)

📈 Research Experience

Research Assistant   **Virginia Commonwealth Univ.**   📅 **Aug 2017 - Present**

- **Extracting Adverse Drug Events (ADE) from Clinical Notes:** Built a contextualized language model-based approach utilizing Bidirectional Encoder Representations from Transformers (BERT). Achieved state-of-art performance for ADE extraction.
- **RelEx: Framework for clinical and chemical relation extraction:** Developed a rule-based approach utilizing co-location information, and a deep learning-based approach utilizing Convolutional Neural Networks (CNNs). Best model achieved a F-measure value of 86%
- **SciREL: A System for Semantic Relation Extraction and Classification:** Developed a feature-vector based system to extract and classify explicit semantic relations. System is trained in the ACL corpus and applied for general English domain.
- **Identifying adverse effects in English tweets for unbalanced data:** Developed a supervised binary classification system to automatically identify the Adverse Effects (AE) in English tweets using Convolutional Neural Networks (CNNs). Experimented techniques to deal with the unbalanced nature of the data and found GloVe trained on twitter is optimal for word representation.

Student Researcher   **Univ. of Peradeniya**   📅 **Jan 2015 - Sept 2015**

- **Extracting bio-medical interactions between drugs from literature using Natural Language Processing (NLP) techniques:** Focuses on extraction of Drug-Drug interactions (DDIs) in biomedical articles from databases such as DrugBank and MedLine. Utilized a feature engineering based approach using NLP techniques. Achieved a encouraging F-measure value of 76.9%.

📈 Industrial Experience

Software Engineer   **MillenniumIT, Sri Lanka**   📅 **Mar 2016 - Jul 2017**

- **Software development in the capital market domain:** Worked in developing (Java) and testing software for the London stock exchange group (LSEG).

Trainee Software Engineer   **IFS R&D, Sri Lanka**   📅 **Sept 2015 - Dec 2015**

- **Mobile app development project using Android:** Travel app custom designed for the use of IFS Employees who travel abroad for business purposes.

📈 Teaching Experience

Teaching Assistant:
- Artificial Intelligence
- Programming languages
- Introduction to Natural language processing (NLP)

Student Supervision

Mentored three undergraduates and a high school student on their research projects.
- **Cora Lewis** (High school student) - Developed a supervised binary classification system to automatically identify the Adverse Effects(AE) in English tweets.
  Publications: NLP@ VCU: Identifying adverse effects in English tweets for unbalanced data.
- **Gabby Gurdin** (Undergraduate student) - Developed a system for Named Entity Recognition (NER) – the automatic identification of chemical reaction parameters from the corresponding text.
  Publications: NLPatVCU CLEF 2020 ChEMU Shared Task System Description.
- **Neha Dil** (Undergraduate student) - Developed a system for extracting experimental design factors for the systematic review process from the method section of biomedical journal articles and extracting relations automatically from text
  Publications: TAC SRIE 2018: Extracting Systematic Review Information with MedaCy.
- **Steele Fransworth** (Undergraduate student) - Developed Pseudobert, a program to automatically generate pseudo training data for relation extraction.

Awards

- **Hackathon:** Won third place in *RamHacks 2019* for designing web app to recommend articles in real-estate domain using deep learning and natural language processing techniques.
- **Nominated for *Best student paper award*:** One of the six finalists nominated for the best student paper award in AMIA-2021.
- **University award:** University award (2015) for excellent performance in academic studies, by the Univ. of Peradeniya (Peradeniya, Sri Lanka).
- **CAPWIC scholarship:** To attend and present in the ACM Capital Region Celebration of Women in Computing (CAPWIC) conference.
- **Travel award:** To present in American Medical Informatics Association (AMIA) - 2019 & 2021 by VCU, USA.

Publications

- **Mahendran, D**. & McInnes, B. (*2021*). Extracting Adverse Drug Events from Clinical Notes. In:Proceedings of the American Medical Informatics Association (AMIA) Summits on Translational Science, March 2021. *[Nominated best student paper]*.
- **Mahendran, D**. & Lewis, C. & McInnes, B. (*2020*). NLP@VCU: Identifying adverse effects in English tweets for unbalanced data. In Proceedings of the Fifth Social Media Mining for Health Applications Workshop Shared Task, 158-160.
- **Mahendran, D**. & Gurdin, G. & Lewinski, N. & Tang, C. & McInnes, B. T. (*2020*). NLPatVCU CLEF 2020 ChEMU Shared Task System Description. In: the Conference and Labs of the Evaluation Forum (CLEF) 2020 Working Notes.
- He, J., Nguyen, D. Q., Akhondi, S. A., Druckenbrodt, C., Thorne, C., Hoessel, R., **Mahendran, D**., Verspoor, K. (2020). An extended overview of the CLEF 2020 ChEMU Lab. In Proceedings of the Conference and Labs of the Evaluation Forum (CLEF) 2020, *2020*.
- **Mahendran, D**. & Brahmana, C. & McInnes, B. (*2018*). SciREL at SemEval-2018 Task 7: A System for Semantic Relation Extraction and Classification. In Proceedings of the 12th International Workshop on Semantic Evaluation, 853-857.

- Mulyar, A. & **Mahendran, D**. & Maffey, L. & Olex, A. & Matteo, G. & Dill, N. & Lewinski, N. & McInnes, B. T. (*2018*). TAC SRIE 2018: Extracting systematic review information with MedaCy., Strain.
- **Mahendran, D**. Nawarathna, R. D. (*2016*). An automated method to extract information in the biomedical literature about interactions between drugs. In 2016 sixteenth international conference on advances in ICT for emerging regions (icter), 155-161. IEEE.

Published Abstracts

- Mulyar, A. & **Mahendran, D** & McInnes, B. T. (*2018*). Medical Information Extraction with MedaCy. In National NLP Clinical Challenges (n2c2).

Presentations

- ***Extracting Adverse Drug Events from Clinical Notes*** - Virtual Informatics Summit, American Medical Informatics Association (AMIA), March 2021.
- ***RelEx: A system for multi-class clinical relation extractio*n** - Student research shorts, ACM Capital Region Celebration of Women in Computing, March 2021.
- ***NLPatVCU CLEF 2020: ChEMU Shared Task System Description*** - Conference proceedings, CLEF 2020 Conference and Labs of the Evaluation Forum, September 2020.
- ***RelEx: A system for clinical relation extraction via Convolutional Neural Network*** - NLP working group, American Medical Informatics Association (AMIA), Nov 2019.

Conference & Jornal Reviewing

- American Medical Informatics Association Annual Symposium (AMIA) - 2021
- CLEF 2020 Conference and Labs of the Evaluation Forum - 2020
- American Medical Informatics Association Annual Symposium (AMIA) - 2020
- Public Library of Science (PLOS) - 2020
- Empirical Methods in Natural Language Processing (EMNLP) - 2020
- American Medical Informatics Association Annual Symposium (AMIA) - 2019
- North American Chapter of the Association for Computational Linguistics (NAACL) - 2019
- International Workshop on Semantic Evaluation (SemEval) - 2018

Volunteer service

- Session chair, American Medical Informatics Association (AMIA) Virtual Informatics Summit - 2021.
- Presenter, Code RVA, VCU - 2018, 2019
- Organizing committee, Human System Interaction (HSI) conference - 2019
- Organizing committee, Cyber robotics coding competition (CRCC), VCU - 2020
- Organizing committee, Code RVA, VCU - 2018, 2019
- Organizing committee, High school coding competition, VCU - 2018, 2019
- Organizing committee, Nano technology day, VCU - 2017, 2018, 2019
- Organizing committee, Computer science (CS) day, VCU - 2018, 2019
- Organizing committee, Computer science (CS) Family Day, VCU - 2019