



2022 Summer Conference on Applied Data Science June 13-August 5, 2022 North Carolina State University, Raleigh, NC

The Laboratory for Analytic Sciences at North Carolina State University is seeking students to join the Summer Conference on Applied Data Science (SCADS), which will take place during eight weeks on campus in Raleigh, NC.

We invite graduate students to apply (advanced undergraduate students will be considered on a case-by-case basis.) Students will be expected to attend in person in Raleigh for the duration of the program and will receive both a stipend and support for travel and lodging.

The overarching, multi-year challenge of SCADS is to generate tailored daily reports for knowledge workers that capture information relevant to their individual objectives and interests. These tailored daily reports, or “TLDRs,” are fairly short and aim to proactively provide individuals with information relevant to their needs, objectives, and interests. The content of a report may be drawn from any number or variety of sources. The TLDRs are auto-generated through development and application of modern AI/ML, user modeling, and information retrieval technologies and methods. The Intelligence Community (IC) manually produces a report of this nature for the U.S. President every morning, summarizing new updates of intelligence information that the President needs to know. Conceptually, AI/ML technology could produce something similar for all IC knowledge workers.

Student applicants should demonstrate research experience (or corresponding coursework) in one or more of the following areas:

Machine Learning and Data Analysis

- Comprehensive knowledge of and experience with applied machine learning concepts
- Experience with foundational models (e.g., BERT, GPT-3, etc.) and transfer learning techniques.
- Experience using deep learning software (e.g. Tensorflow, Pytorch, Theano, Keras, etc.) and architectures (e.g. Encoder-Decoder, LSTM, RNNs, Transformers, etc.)
- Experience working on natural language processing (NLP) challenges, specifically around text summarization, relationship extraction, and question answering
- Experience with recommendation systems
- Understanding of statistical and mathematical methods used in model training / evaluation, and hypothesis testing.
- Experience analyzing user interaction logs
- Experience with information extraction pipelines for text, video, and / or audio

Knowledge Management

- Experience working with graph databases and maintaining operational knowledge graphs
- Experience with building scalable knowledge-based information retrieval systems
- Experience building comprehensive, intuitive, semantic ontologies representing knowledge across multi-source, multi-modal data
- Experience evaluating knowledge management systems

Software Engineering

- Experience setting up and using computing environments using Amazon Web Services (AWS), including S3, EC2, EMR, etc.) to design, develop, deploy, maintain, and monitor web applications within AWS infrastructures
- Experience with data science tools including Python scripting, numpy, scipy, pandas, matplotlib, scikit-learn, and jupyter notebooks, and bash scripting
- Proficiency with Python programming, including numeric, data analysis, scientific, and machine learning libraries such as NumPy, pandas, scikit-learn, and/or PyTorch
- Programming web applications in HTML, CSS, and JavaScript using React JS
- Experience implementing and using RESTful web services
- Experience using cloud-based machine learning services, such as AWS and Azure Machine Learning services

This position is full-time (40 hours per week) for the duration of SCADS. To apply, students will need to submit a resume/CV, cover letter, and contact information for three references via the [SCADS Student Intern Application](https://ncsu-las.org/2022-scads) linked from our website, <https://ncsu-las.org/2022-scads>. The priority deadline for applications is 11:59 p.m. EST on February 18th. Students selected for an interview will be notified by February 25th. The application will remain open until March 4th and secondary interview offers will be extended by March 11th.