

Texas Undergraduate Research Day

Transforming Texas Through Undergraduate Research

Two CSE students, John Forrest and Vladimir Jovanovic, represented SMU at the first Texas Undergraduate Research Day held on February 14 at the state capitol building. This first, to be annual, statewide event had 84 poster presentations involving over 140 undergraduate students from 51 different colleges and universities across the state. The purpose of the event was “to showcase the experiences of undergraduate students engaged in research for Texas legislators and the public through high-quality poster presentations.” During the day, state legislators and other officials visited the poster session. John and Vladimir proved to be excellent ambassadors of SMU and CSE by providing excellent descriptions of their research to the visitors. Both John and Vladimir have been supported by an NSF REU grant awarded to Professors Dunham and Hahsler.



John Forrest is a senior majoring in Computer Science and will graduate in May with the first ever CSE Departmental Distinction granted by the department. He has accepted a position at Microsoft after graduation. John



is working with his advisor, Professor Michael Hahsler, in the area of data stream clustering. Data streams constantly produce typically high-dimensional data which needs to be processed in real-time. Many modern applications generate such data streams (e.g., web click-stream data, computer network monitoring data, telecommunication connection data, readings from sensor nets, stock quotes, etc.). The title of John’s research is “ClusterDS: A Framework For Clustering Data Streams in R.” R is a free programming environment widely used for data analysis and data mining. Although a commercial high-performance version of R is available, currently no standard way exists to process data streams. By creating a

flexible framework to represent data streams and data stream mining algorithms, John's work contributes significantly to the data mining community.

Vladimir is a junior majoring in both Computer Science and Psychology. Vladimir has many research interests not only in Computer Science. He has also been supported by an SMU URA award from Professor Bob Kehoe in the Physics Department. He is also currently working on a research project through the Innovation Competition at the Caruth Institute for Engineering Education. Vladimir is working in Bioinformatics under the direction of advisor Professor Margaret Dunham. The title of Vladimir’s research is “Unique Base Pair Count Vectors in Analysis of Nucleotide Sequences.” Secondary and tertiary structures in single-stranded nucleotide sequences are important for DNA and RNA sequences when binding to particular molecules. Vladimir has developed a novel method of base pair counting that can take into account these structures. The base pair counter looks at complementary base pairings of a folding nucleotide sequence. This method thus creates a unique vector for each nucleotide sequence that leads to a simple but effective summarization of the possible secondary and tertiary structures that the sequence can have. The method can be used in a variety of applications in genetics and bioinformatics and has been shown to be useful in at least one case of aptamer classification. Specifically, using the novel method with data mining classifiers, a truth positive rate of 94.4% to 100% was achieved depending on the classifier used.



For more info on the Intelligent Data Analysis Lab IDA@SMU see: <http://lyle.smu.edu/IDA>