Assessing the Value of Consumer-Provided Big Data

Online consumer reviews are one of the most important drivers to encourage potential customers to buy products or services. Because of this, businesses are striving to find the best practices to utilize this information. However, consumers have many different options when they provide their opinions online. More than just numerical ratings, rapidly-advancing mobile platforms now promote various forms of online interactions between businesses and consumers. Online consumers leave behind trails regarding their product or service experiences. This generates a tremendous amount of data in unstructured formats such as texts, image, voice, and video—so-called “Big Data.” Deep Learning can be used to efficiently mine unstructured textual or image data on a large scale through customized neural network methods and computing systems. Young Jin Lee and his team are working to find the ways to exploit deep learning to create business values for firms. The team collects online consumer reviews and feedbacks in texts and photos from publicly-available data sources provided by companies such as Yelp and Airbnb. The data is processed by deep learning to extract and create new features and put them into an econometrics framework that evaluates how online consumer opinions in texts and images can affect businesses and peer consumers. This research will offer firms with advanced and practical social media strategies specifically-tailored information for more personalized products and services to meet various needs from consumers.

FOR MORE INFORMATION

Associate Professor Young Jin Lee
Mailing Address:
Daniels College of Business, Suite 580
2101 S. University Blvd.
Denver, CO 80210
Email: YoungJin.Lee@du.edu
In today’s rapidly changing technological, economic, and geopolitical world, the students of today must develop a complex set of cognitive, motivational, and socio-emotional attributes—termed 21st Century Skills. These skills will allow them to flexibly adapt to the challenges of the future. Because of this, educators around the world are now emphasizing these higher-order skills in their teaching and downgrading the importance of rote memorization. Unfortunately, formalized educational measurement and testing has not kept pace with this shift. Today, most statistical models used to validate educational measures are not able to incorporate open-ended or unstructured data sources such as student writing, drawings, or speech.

Current attempts to circumvent this issue in large-scale testing typically involve hiring graders to rate student responses, a process that is both highly expensive and largely error-prone. Recent advances in artificial intelligence may provide solutions. By integrating state-of-the-art natural language processing methods with psychometric models, open-ended measures may be validated for use, revolutionizing the science and practice of educational assessment through artificial intelligence.

This interdisciplinary program of research, sometimes termed computational psychometrics, represents a major area for innovation that has the potential to reshape educational assessment—from major high-stakes standardized tests, to smaller-scale measurement in schools. With the full development of this interdisciplinary field, educational assessments will no longer require close-ended data sources (e.g., multiple-choice items), and will become much richer sources of information about students. This educationally- and psychologically-oriented application of advanced artificial intelligence models has the potential for massive societal impact in the way student learning processes and trajectories are understood both by social scientists as well as educational practitioners and policymakers.

Additionally, the Massive Texts Lab, led by Dr. Peter Organisciak, develops computational methods to learn about cultural, historical, and linguistic trends from digital libraries. An algorithm does not have the same intuitive understanding of what words mean as people have when they read. To help a system understand a text, the Massive Texts Lab needs to train deep learning models that can represent documents not as a set of words, but as a mix of concepts that those words represent. From that representation, the Massive Texts Lab can better analyze documents.

Algorithms try to honestly reflect the world presented to them, and the models they train end up reflecting the trends or prejudices of the texts that they learned from. With historical collections, this provides a lens to see how biases change across time, genres, or countries of publication. One current IMLS-funded project is analyzing relationships between books at scale—identifying variants and alternate versions of the works and recommending thematically similar books. This work will help libraries improve the findability of materials in their collections.

FOR MORE INFORMATION

Assistant Professor Denis Dumas
Mailing Address: Ruffato Hall, Room 233
1999 E. Evans Ave., Denver, CO 80210
Email: Denis.Dumas@du.edu

Assistant Professor Peter Organisciak
Mailing Address: Ruffato Hall, Room 247
1999 E. Evans Ave., Denver, CO 80210
Email: Peter.Organisciak@du.edu
Making Big Data Exploration Affordable

Massive georeferenced data are produced daily. From on-vehicle sensors which generate GPS tracking data, to climate models which provide terabytes of data describing the weather conditions (e.g., pressure, temperature), the availability of big data brings tremendous opportunities to examine various social and environmental phenomena. With the knowledge generated from big data, we are able to make more informed decisions.

However, processing massive geospatial data is computationally intensive and often exceeds the processing capabilities of conventional computers. Jing Li’s team focuses on leveraging GPU (Graphical Processing Unit) techniques and cloud computing services to process massive geospatial data efficiently. The team built GPU-based computing algorithms that can process multiple datasets concurrently with many GPU cores within desktop or laptop environments, making high-performance computing capabilities affordable. The algorithms can also run within the cloud environments where multiple machines are available to meet the needs of processing even larger datasets.

The outcomes of this work go beyond building parallel algorithms that empower researchers and scientists to conduct big data-driven research. In many cases, the time and resource costs of big data processing are unpredictable. The team has also developed an intelligent algorithm that can assess the performances of computing resources in processing historical data and predict the workloads of processing current data with an online-learning model.

FOR MORE INFORMATION

Associate Professor Jing Li
Mailing Address:
Boettcher Center West, Room 106
2050 E. Iliff Ave., Denver, CO  80210
Email: Jing.Li145@du.edu

3D visualization frameworks simulate the evolution of extreme weather phenomena such as dust storms. With this framework, scientists can validate the accuracy of complex quantitative models efficiently. The algorithms usually run 10x to 100x faster compared to serial processing.

2018 Another RECORD BREAKING Year

Principal Investigators were funded by external awards. One out of every four faculty members on campus has an active external funded research award. 115 new contracts were awarded.

In research and development expenditures. This helped support more than 134 graduate students, who received more than $3.9 million in stipends, tuition waivers & health benefits.

Growth in research expenditures from FY17, representing a total growth of more than 40% in three years. The number of external funded faculty also increased 7% from FY17.
Every year, up to two million youths in the U.S. will experience homelessness, and estimates suggest between 39 and 70 percent of these young people abuse drugs or alcohol. Group-based intervention programs offer promising means of preventing and reducing substance abuse by encouraging homeless youth to share their experiences, learn positive coping strategies, and build healthy social networks. While effective, unfortunately, inappropriate intervention groups can result in an increase in deviant behaviors among participants, a process known as deviancy training. Deviancy training occurs when high-risk youth are grouped together and reinforce negative behaviors and attitudes.

To address the issue of deviancy training, Assistant Professor Anamika Barman-Adhikari along with researchers from the University of Southern California’s Center for Artificial Intelligence in Society (USC CAIS) developed an AI-based decision aid, called GUIDE (GroUp-based Intervention DEcision aid). GUIDE assists interventionists in substance abuse prevention by giving recommendations regarding how to structure intervention groups for substance abuse prevention so that deviancy training can be minimized. Group-based interventions have typically placed participants into intervention groups randomly.

In contrast, GUIDE used data collected from young people experiencing homelessness in Los Angeles to build an algorithm that takes into account both how the individuals in a subgroup are connected -- their social ties -- and their prior history of substance abuse to structure the intervention groups.

When the team tested the algorithm against the data, the simulation results were remarkable: compared to randomly-assigned groups, deviancy training was reduced by almost 60 percent in AI-assigned groups. The study, called Influence Maximization for Social Network Based Substance Abuse Prevention, was published in the AAAI conference on Artificial Intelligence. The researchers are continuing to work with Urban Peak, and plan to deploy the tool to optimize intervention group strategies for young people experiencing homelessness in Denver in fall 2019.
Comparing Neural Network Architectures for Financial Time Series Forecasting

Associate Professor Matthew Rutherford and MS computer science candidate Nathan Egan are studying the applications of artificial neural networks (ANNs). In general terms, a neural network is a collection of computational elements that starts with input data, uses math to transform that data, and then passes its result along to other elements in the network to create output data. The network ‘learns’ how to map inputs to outputs through a process of systematically changing the parameters the network uses to pass information around.

Using existing research from leaders in the field, the team can focus on applying these neural networks in new and unique domains such as financial and economic time series forecasting. They are examining how to apply neural networks, including those that have been designed for computer vision tasks, in the context of economic and financial time series forecasting. The results of this research will provide a more comprehensive understanding of how ANNs can be applied across domains and will help data scientists evaluate the tools they can use to solve problems.

FOR MORE INFORMATION
Associate Professor Matthew Rutherford
Mailing Address: Engineering & Computer Science,
Room 345
2155 E. Wesley Ave. Denver, CO 80210
Email: Matthew.Rutherford@du.edu

Nathan Egan
MS Computer Science Candidate
Email: Nathan.Egan@du.edu

RECENT GRANTS AWARDED

Director Charmaine Brittain (GSSW) $15K & 19K
Grant from Coordinated Care Services, Inc. for "ACCESS Training" & Grant from the Wyomong Department of Family Services for "WY Cheyenne Leadership Support"

Chair Anne DePrince (AHSS) $130K
Grant from the Rocky Mountain Victim Law Center, subcontract from the Colorado Division of Criminal Justice for "Legal Information Network of Colorado (LINC) Expansion Project"

Faculty Elizabeth Escobedo, Faculty Carol Helstosky & Assistant Professor Esteban Gomez (AHSS) $225K
Grant from the U.S. Department of Veterans Affairs for "NCA’s Veterans Legacy Program (VLP)"

Assistant Professor Donald Gerke (GSSW) $7K
Grant from Washington University St. Louis, subaward from SAMHSA for "SPNS Subcontract"

Director Lotta Granholm-Bentley (KHIA) $376K
Grant from the University of Kentucky, subaward from the National Institutes of Health, for "MTOR activation and the pathogenesis of Alzheimer’s Disease in Down Syndrome"

Director Lotta Granholm-Bentley (KHIA) & Associate Professor (NSM) $3.3M
Grant from the University of Kentucky, subaward from the National Institutes of Health, for "MTOR activation and the pathogenesis of Alzheimer’s Disease in Down Syndrome"

Executive Director Kendra Whitlock Ingram (Newman) $20K
Grant from the National Endowment for the Arts for "Newman Center Presents"

Suzanne Kerns (GSSW) $2.29M
Grant from the Colorado Office of State Planning for "Pay for Success - MST"

Assistant Professor Jonathan Moyer (JKSIS) $75K & 65K
Grant from the African Development Bank for "SDG3 Analysis for African Using International Futures" & Grant from the United Nations Development Program for "Country Study on the Impact of War on Development in Yemen"

Assistant Professor Christine Nelson (MCE) $173K
Grant from the University of New Mexico, subaward from W.K. Kellogg Foundation for "New Mexico Learning and Education Consortium (NMLEC)"

Librarian Kim Pham, Beck Curator Jeanne Abrams, Librarian Kevin Clair, & Associate Dean Jack Maness $41K
Grant from the University of Nevada, Las Vegas, subaward from the Andrew Mellon Foundation, for "Uncovering Health History: Opening TV patient records in the early 20s as data"

Senior Research Associate Amy Roberts (Butler) $50K
Grant from the University of Nebraska for "EduCare"

Adjunct Professor Maria Vukovich (GSSP) $22K
Grant from the University of Minnesota for "University of Minnesota Partnership"

Co-Director Sarah Watamura (AHSS) $100K & $52K
Grant from the Colorado Department of Human Services for "Implementing the Seedlings Curriculum to Build Parental Resilience" & Grant from the Colorado Department of Human Services for "Roots Workshops and Coaching"
Social Network-Driven Algorithms May Help Students In Crisis

Suicide is a significant public health issue among college students in the United States. One out of every ten college students reported suicidal thoughts in the past year alone. Yet, the vast majority of college students who experience suicidal thoughts do not seek help from trained professionals.

Gatekeeper training—teaching people to recognize and support someone in an emotional or mental health crisis—is often used to encourage help-seeking behavior among college students. However, the effectiveness of a gatekeeper program may depend on which people are recruited for training.

To date, the traditional gatekeeper approach has involved training a group of people who are in institutional roles that reach a large segment of the student body (e.g., Resident Assistants). However, two major problems exist with that approach. First, students talk more to their friends about suicide than anyone else, that friendship information is not used during recruitment for training. Second, not everyone who attends training will perform equally well as gatekeepers—also not accounted for during training recruitment.

As a result, we don’t know how many students are in a relationship with someone who they would talk to about suicide and has been effectively trained (i.e., true degree of network coverage).

In last 18 months, we’ve been designing a new strategy for gatekeeper recruitment. This strategy relies on two components: (1) social network methodology to map student friendships and (2) an algorithm that can guide recruitment to maximize coverage of the student network while accounting for uncertainty in gatekeeper performance.

Our next step is to compare the traditional gatekeeper approach to recruitment with our network-driven algorithmic approach. Our main goal is to increase the likelihood that students experiencing a crisis get the support they need.