

MAYANK NAGPAL

Phone: +1 224-704-5536 | Email: m.nagpal@northeastern.edu

EDUCATION

The Pennsylvania State University

State College, PA, USA

Ph.D. Business Administration

2015 – 2021

- Thesis: “Essays on Textual Analysis in Marketing & Management: Uncovering Underlying Intent from Words”
- Dissertation Committee: J. Andrew Petersen (Penn State University), Arvind Rangaswamy (Penn State University), Gary Lilien (Penn State University), Mark Desjardine (Penn State University)
- Research emphasis:
 - Substantive: Search Engine Marketing, Digital Marketing, Strategic Marketing Management, Scale Development
 - Methodological: Text Analysis, Causal Research / Natural Experiments, Machine Learning / Big Data

Indira Gandhi Institute of Development Research

Mumbai, India

MSc. Economics

2008 – 2010

- Research emphasis: Econometric research techniques

Hindu College, University of Delhi

New Delhi, India

BSc. Honors Mathematics

2005 – 2008

Publications

-
- Nagpal, Mayank, J. Andrew Petersen, Keyword Selection Strategies in Search Engine Optimization: How Relevant is Relevance?, Journal of Retailing, 2020, ISSN 0022-4359, <https://doi.org/10.1016/j.jretai.2020.12.002>
 - Created a new data-driven framework to select keywords, based on factors such as brand equity and customer search intent, to drive improved Search Engine Optimization
 - Honors: Recipient of the Marketing Science Institute, Young Scholars Grant Proposal

Working Papers

-
- ‘From Words to Indices: Measuring Digitalization from Corporate Reports’ Mayank Nagpal, J. Andrew Petersen and Arvind Rangaswamy, Working Paper, Penn State University
 - Developed a novel approach to measure latent strategic constructs such as Market, Customer or Digital Orientation using text from annual 10k reports using Natural Language Processing (NLP) methods
 - Used Supervised Latent Dirichlet Allocation (sLDA) model to estimate the digital weights of words/phrases.
 - Operationalized and validated the method for measuring Digitalization against existing indices independently by Gartner and Mckinsey research and against investment in digital applications using text from annual 10-K reports across 20 years
 - Invited for presentation at the European Marketing Academy (EMAC) Conference, 2019 (Hamburg, Germany)
 - In preparation for submission to the Journal of Marketing
 - Currently working on a paper to study the financial impact of digital initiatives as a response to the disruption caused due to the Covid-19 pandemic (Co-Authors: Dr. Arvind Rangaswamy, Professor of Marketing at Penn State University and Dr. Ashish Galande, Assistant Professor of Marketing, Indian Institute of Management, Udaipur).
 - Using NLP based network analysis to study the digital transformation as response to the Covid-19 pandemic
 - Analyzed unstructured text data generated from analyst calls, 10Q reports and press releases.
 - Analyzing the difference in responses of firms based on the varying degree of existing digitalization.
 - Analyzing the impact of digital initiatives across industries allowing varying degrees of random mixing of people.

TEACHING EXPERIENCE

Pennsylvania State University

State College, PA, USA

Teaching Instructor

2018-2021

- Designed and taught graduate and undergraduate courses in Marketing and Analytics (Avg. Evaluation: 5.6/7)
 - Marketing Research: Undergraduate course in Summer 2018, Fall 2018, Spring 2020, Fall 2021
 - Programming Skills for Business Analytics: Graduate course in Fall 2020
- Teaching Assistant: Graduate Course on Data Mining for Business (Spring 2021)

CONFERENCE PRESENTATIONS

-
- 2019 Journal of Retailing Thought Leadership Conference on Metrics and Analytics in Retailing, Atlanta, GA

- ‘Keyword Selection Strategies in Search Engine Optimization: How Relevant Is Relevance?’, Mayank Nagpal and J. Andrew Petersen
- 2019 Haring Symposium, Indiana University, Bloomington, IN
 - Discussant: ‘Studying the impact of Advertising Clutter on Sales’, Sean Melessa and Paul Hoban
- 2018 INFORMS Marketing Science Conference, Philadelphia, PA
 - Presenter: ‘Keyword Selection Strategies in Search Engine Optimization: How Relevant Is Relevance?’, Mayank Nagpal and J. Andrew Petersen

PROFESSIONAL EXPERIENCE

D'Amore-McKim School of Business, Northeastern University **Boston, MA , USA**
Visiting Assistant Professor of Marketing 2021- 2022

- Teaching Marketing Analytics courses at both undergraduate and graduate level in the academic year 2021-2022.

CRISIL Global Research & Analytics **London, UK**
Senior Consultant 2013 –2015

- Led a team of 5 analysts to design a stress test Scenario Expansion Prototype for all asset classes of a major UK investment bank including Interest Rates, Foreign Exchange, Credit, Equities, and Commodities, and over 150,000 other detailed risk factors.
- The testing prototype enhanced the bank’s stress testing capabilities and ability to meet its regulatory requirements.
- Designed, implemented, and managed analytics-based consulting assignments providing risk management solutions.
- Calibrated enterprise-wide suite of Probability of Default Corporate Rating Models for one of the largest Saudi Arabian banks. The model facilitated the implementation of the Basel II regulatory norms.
- Developed an automated credit approval strategy for one of the largest public sector banks in India.

Australia & New Zealand (ANZ) Bank **Bengaluru, India**
Retail Modelling Analyst 2010 –2013

- Integrated and validated 11 credit risk models including personal loans, credit cards, auto loans and mortgages into a single management system as a part of the merger of National Bank of New Zealand and ANZ New Zealand.
- Designed a model performance monitoring framework for regular monitoring of ANZ Australia’s mortgage portfolio in order to automatically alert the managers when model performance deteriorates.

Indian Institute of Management (Centre for Public Policy) **Bengaluru, India**
Research Associate (Internship) May 2009 – Jul 2009

- Studied the efficiency of the commodity markets under the guidance of Dr. Gopal Naik, Professor at the Indian Institute of Management, Bengaluru writing a research paper titled ‘Efficiency of the Commodity Market - A Co-integration approach’
- Paper invited for presentation at Pondicherry Conference on “Indian Commodity Market Derivatives and Risk Management – The Road Ahead”, 2009.

GRADUATE COURSEWORK

Marketing and Management Courses

- Organizational Research Design, Consumer Behavior, Research Methods in Marketing, Marketing Models, Marketing Management

Economics and Statistics Courses

- Regression Analysis, Econometrics I, Empirical Industrial Organization, Applied Microeconomics Theory, Casual Inference Methods, Empirical Methods in Economics, Development Economics, International Trade, Time Series Modeling for Macroeconomics

EXPERTISE

- **Domain Expertise:** Marketing Strategy, Digital Marketing, Content Marketing, Risk Management, Business Analytics
- **Tools:** R, STATA, SPSS, SAS, MS Office
- **Technical Expertise:** Text Analysis, Probabilistic Modeling, Causal Research, Instrumental Variable, Tobit Modelling, Conditional Mixed Processing (CMP), Natural Language Processing (NLP), Supervised Latent Dirichlet Allocation (sLDA), Latent Dirichlet Allocation (LDA), Structural Topic Modeling (STM)

PAPER ABSTRACTS

- ‘Keyword Selection Strategies in Search Engine Optimization: How Relevant Is Relevance?’ Mayank Nagpal and J. Andrew Petersen, *Journal of Retailing*, 2020, ISSN 0022-4359
 - We build an empirical framework using search queries and organic click data which provides model-based guidance to SEO practitioners for keyword selection and web content creation. Specifically, we study how search characteristics (search query popularity, search query competition, search query specificity, and search intent) and website characteristics (content relevance and online authority) interact to affect the expected organic clicks as well as the organic rank a website receives from the search engine result page (SERP). It is often thought that content relevance is a key factor to improve the effectiveness of SEO. We find, however, that content relevance is an important factor in driving organic clicks only when the consumer is farther along in the customer journey and searching for ways to purchase a product. Whereas, when the customer is at the awareness stage and looking for product information, online authority is the key driver of organic clicks.
- ‘From Words to Indices: Measuring Digitalization from Corporate Reports’ Mayank Nagpal, J. Andrew Petersen and Arvind Rangaswamy
 - Computer Aided Text Analysis (CATA) of firm generated text is increasingly being seen as a less resource intensive alternative to survey data-based methods to measure strategic constructs. However, dictionary based CATA methods are not easy to replicate as they require intensive manual effort from domain experts. To remedy this, we develop a readily replicable method for measuring Strategic Orientations (SOs) such as Market Orientation, Customer Orientation, or Degree of Digitalization of a firm using textual analysis of Annual 10-K reports of firms. The method uses publicly available documents to measure SO scores across firms and across time periods. We demonstrate the validity and value of our approach by operationalizing it for measuring the construct of digitalization. The digitalization score is designed to capture an enterprise-wide view encompassing all aspects of the digital transformation associated with a business, such as automation of business processes through technology, digital innovation, and enhancement of the digital capabilities of a business. Using Parts of Speech (PoS) tagging, Supervised Latent Dirichlet Allocation (sLDA), and a careful review of academic and practitioner research related to digital transformation of businesses, we compile a dictionary of terms (words/bigrams) that reflects the strategic intent of senior management of firms regarding digitalization. The estimates of the sLDA model can be used to assign a SO weight to each term in the dictionary, based on its probability of being used in a set of SO-focused articles. We use the predicted scores from the sLDA model on a firm’s annual 10-K report as a measure of digitalization of that firm in any given year. We validate our measure by comparing the predicted scores from our model for a set of 247 firms against existing measures developed independently by Forbes and Gartner Inc. The validation shows significant predictive strength of the model by achieving a rank correlation of 0.36 and 0.32 between the proposed and existing scores from Gartner and Forbes respectively, which is about 15% and 37% below the target correlation calculated by comparing the correlation against that estimated using a sLDA model directly trained on the validation data.