UConn School of Business

MSBAPM NEWSLETTER

May 2016

UConn MSBAPM congratulates the May-2016 graduates

UCONN
SCHOOL OF BUSINESS

CONTENTS

THIS MONTH

Natural Language Processing (NLP) and Text Analytics – the power of unstructured data

We all know the 3 Vs that eternally define Big Data - Velocity, Volume and Variety. It's not a cliché that Big Data is really the future with tremendous amount of untapped potential. Do you know that more than 80% of all the data in the world is unstructured? Companies collect massive amounts of documents, emails, social media and other text based information to know their customers better, offer customized services, or comply with federal/regulatory authorities. It is intriguing to know that most of this data is unused and untouched. Text analytics, through the use of natural language processing (NLP), holds the key to unlocking the boundless potential of unstructured data. Through this article we will learn the basics of NLP and text analytics, understand some of the use cases where these concepts can be applied and finally learn some of the techniques involved in an implementation of NLP.

Text analytics is the process of analyzing unstructured text, extracting relevant information, and transforming it into useful business intelligence. Text analytics processes can be performed manually, but the amount of text-based data available to companies today makes it increasingly important to use intelligent, automated solutions. It is a broad term that describes tasks from annotating text sources with meta-information such as people and places mentioned in the text to a wide range of models about the documents (e.g., sentiment analysis, text clustering, and categorization). Natural Language Processing (NLP) is a scientific discipline concerned with making natural language, as spoken and used by human beings, accessible to machines. NLP addresses tasks such as identifying sentence boundaries in documents, extracting relationships from documents, and searching and retrieving of documents, among others. NLP is necessarily a means to facilitate text analytics by establishing structure into unstructured text to enable further analysis.

The applications of text analytics and NLP are widespread and spans across multiple industries and business processes. Some of the applications (not an exhaustive list) per each industry are listed below:

Industry	Applications
Finance	Search; Compliance; Entity Matching; Call Center analytics; risk management; Anti-money laundering
Insurance	Sentiment of customer interactions; Problem topics identification; Claim adjuster notes extraction; underwriting analytics
Media and Technology	Social Media analytics; Audio/Video broadcast content analysis; Image Processing
Retail	Brand/Product analytics based on customer feedback; Accounts payable and accounts receivable analytics
Process Control	Problem/Cause augmentation from operator notes
Energy	Price and demand forecasting
Oil and Gas	Operator comment analysis; Drilling efficiency
Legal	Search; Relationship extraction (e.g. X sued Y); Document clustering (clustering documents for similar complaints for class action suit discovery)
Health Care	Medical record content extraction; Drug interaction discovery; Disease outbreak monitoring and control from social media data
Security	Network log analytics; Fraud and hacking detection
Government	Disaster scoping and damage assessment from social media data; citizens' sentiment analysis to a new policy introduction

As we learn about some of the many applications of NLP and text analytics, it is imperative to understand some of the common techniques that majority of NLP

algorithms use in order to successfully implement the text analytics solutions on a given document (unstructured data). Some of the most commonly used techniques, along with an illustration are summarized in the table below:

Reference: https://blog.pivotal.io/channels/big-data-pivotal

Technique	Definition	Unstructured Data	Output of the technique
Sentence Segmentation	Identifies sentence boundaries. Punctuation often marks sentence boundaries, there could be many exceptions	Frank met the president. He said: "Hi! What's up - Mr. President?	**Sentence 1:** Frank met the president. **Sentence 2:** He said: "Hi! What's up - Mr. President?
Tokenization	Identifies word boundaries. Individual words, numbers and other single coherent constructs	My phone tries to change 'eating' to 'dating'. #hateautocorrect	[My] [phone] [tries] [to] [change] ['] [eating] ['] [to] ['] [dating] ['] [.] [#hateautocorrect]
Stemming/Lemmatization	Stemming strips the 'ending' of words.	eating, ate, eat	eat, eat, eat
Part-of-Speech tagging	Assigns each word in a sentence its respective part of speech, such as Verb, Noun, Adjective etc.	If you build it, he will come	If you build it, he will come (IN) (PRP) (VB) (PRP) (PR) (MD) (VB)
Parsing	Derives the syntactic structure of a sentence. It is often used as a prerequisite for 'Entity Recognition' technique. In the example, 'John and Frank' are to be treated as conjunctive noun phrase (NP) and both of them were involved in an action 'went' (VP)	John and Frank went into a bar	(S (NP (NP John) and (NP Frank)) (VP went (PP into (NP a bar))))
Named Entity Recognition	Identifies entities such as persons, locations, and times within a document	Let's meet John in DC at 6 pm	Let's meet **John** in **DC** at **6 pm**. Person Place Time
Co-reference resolution	Attempts to identify multiple mentions of an entity in a sentence or document and marks them as the same instance	John drank a beer. He thought it was warm	**John** drank a **beer**. **He** thought **it** was warm

Market Basket Analysis – A Sneak Peak

Market Basket Analysis is referred to as the rule's antecedent condition of A and B is the consequent. The rule is interpreted as "If A occurs in the market basket, then B also occurs in the market basket."

The support and confidence for rule AB are defined as: If A and B occur together in at least X% of the market baskets, then the support for this rule is X.

Of all market baskets containing A, if at least X% also contains B, then the confidence for this rule is X.

CONVERTING RULES INTO INPUTS

Generally, MBA tools don't automatically create features or dummy variables to identify individuals (i.e., the values of the transactional-id variable) having a particular rule in their market basket, for the number of possible rules is typically too large. For example, suppose there are 1000 possible items (i.e., web-pages) that a surfer can visit or put into their market basket. Then there are 1000*999/2 = 499,500 distinct possible combinations of items taken two at a time, many of which MBA will identify as rules.

Instead of creating a dummy variable corresponding to a particular rule of interest, for example, the analyst might create a dummy variable based upon a set of rules meeting some minimum value of confidence, support, or lift.

For the purpose of using confidence as a criterion for creating RB inputs, the question as to how transactional categorical data can be collapsed into a single row per person or transaction-id value arises, and a method for transforming transactional tables into "modeling" tables is of relevance.

id	Item	Count	Amount	Banner-ad
10019	/hr/cafe/	1	0	0
10019	/hr/cafe/deli/	1	0	0
10019	/hr/cafe/forms/	1	0	0
10019	/hr/cafe/recipes/	1	$5.65	1
10019	/hr/cafe/recipes/Soup	1	0	0
10019	/hr/vendor.html	1	0	0

Here the person "10019" placed these six items in their market basket; that is, this person visited the pages represented by the corresponding directories (each directory references an index.html file). The Count, Amount, and Banner-Clicked columns for this person can be collapsed into the following row using the sum statistic:

Id	Total Amount	Page Count	Banner-ad
10019	$5.65	6	1

Note, however, that by summarizing in this fashion information is lost. Here the specific item information or list of items placed in the market basket is lost, illustrating that the real challenge in collapsing transactional data lies in dealing with the categorical rather than the continuous data.

TRANSPOSITION OR THE TRANSPOSE METHOD

Transposition is one approach to preserve the specific item information. Using the transpose procedure in SAS, the resulting row of the "modeling" table corresponding to the previous transactional data has the id column and six new columns, each corresponding to a distinct item and containing the value one. Here the count is used, but the sum could likewise have been chosen, in which case all columns would take a value of 0 except the "/hr./cafe/recipes/" column, which would take the value $5.65. Perhaps the analyst wants both the total amount and total count of each item placed in the market basket; in this case the modeling table would have 13 new columns.

USING CONFIDENCE AS A CRITERION FOR A RECLASSIFICATION SCHEME

For the purposes of converting a transactional table that has a categorical column with "too" many levels into a modeling table, the general approach is to reduce the number of levels of the categorical variable

using a reclassification scheme and use transposition to generate the candidate inputs. The confidence statistic can be implemented as a reclassification scheme by modifying the where clause used in the previous SQL code to select all rules having a specified minimum value of confidence.

One obvious criterion for reducing the dimensionality of a categorical variable having "too" many levels is to collapse all levels that correspond to a rule having a value of confidence of 100%, for such combinations occur together in every situation and can therefore be considered as a single item.

Consider, for example, the legendary rule, diapers with beer (or D to B for short). And suppose that every time diapers are purchased beer is purchased. Then the rule D to B would have a 100% confidence value, and it seems that the two levels "diapers" and "beer" could justifiably be collapsed into the new single level labeled "diapersbeer." By replacing the two previous levels, "diapers" and "beer", with the one new level "diapersbeer", it seems that the analyst reduces the number of levels of the categorical variable by one.

USING SUPPORT OR LIFT AS A CRITERION FOR CREATING A RB (RULE BASED) DUMMY VARIABLE

As an alternative to choosing some particular rule of interest as the criterion for creating a RB dummy variable, the analyst might choose people based upon whether they have any rule in their market basket with a value of support greater than or equal to, say, 50, which is chosen solely for the purpose of illustration. The path that the interpretation of such a dummy variable might take would be something similar to the following. If the support for the rule is 50, then the rule occurs in 50% of the market baskets. If in addition the confidence is 100% for both this rule and it's opposite, then a dummy variable created using this rule would be expected to have half the values equal one and half the values would equal zero.

Suppose, for example, fraudulent drug claims are being modeled, and that two of the columns in the claims table are id and drug. If a person claims two drugs, then the person has two rows in the claims table. One might then create potential new features. It is conceivable, for example, that a rule is produced that corresponds to what is often referred to as the AIDS cocktail or combination of drugs, for this cocktail has become a common treatment. Indeed, to the extent that doctors on the whole treat various diagnoses with the same drug combinations, we would expect the rules output to identify common drug treatments. If fraudulent behavior is a typical, one might create a fraud predictor (dummy) variable based upon all rules having a level of support greater than or equal to, say, 5, for such a variable would serve to identify those individuals in the claims table having one or more of these rules in their market basket. This feature would then indicate typical behavior or drug combinations. And if the fraud predictor proved to be significant, the analyst would certainly want to add the newly discovered input to the previously used or "traditional" modeling table and develop a new and improved fraud detection model.

CONCLUSION

One of the primary objectives in data mining is to develop predictive models and improve the accuracy of existing predictive models. And one of the essential challenges toward this end lies in the discovery of new features or inputs to foster predictive accuracy.

Reference:
http://robotics.stanford.edu/~ronnyk/WEBKDD2000/papers/gayle.pdf

How to Use Your Summer Break for Career Development

By Katherine Duncan

Congrats, you did it! Another semester is in the books and now you will get some well-deserved time off from school. However, during your downtime it is best to refocus on your career development goals! Here are some ideas and tips for using your vacation time wisely!

- Attend holiday parties for networking – use the holidays as an ice breaker to start conversations (Memorial Day, 4th of July, Summer BBQs, Labor Day)
- When you get together with friends/family talk to them about their professions and companies. What trends are they seeing in the market? How do they like where they work?
- Do research on target companies, make a list up to 40+ that you want to pay close attention to
- Work on your LinkedIn profile, add connections, participate in group discussions
- Volunteer! It's good for networking, adding to your skills, and helping others
- Ask people in your network for informational interviews, you have the time and flexibility!
- Commit to your future job search by preparing your application documents – resume & cover letter

Don't take the summer off, utilize your time to get ahead! The next semester, classes, and homework will be here before you know it. So plan ahead to be more successful in the long run!

"Blue Agent" – My First Week Intern Experience

By Srinivas Godavarthi – Student MSBAPM

"Life is change. Growth is optional. Choose wisely." - Karen Kaiser Clark

This is my first internship experience ever and I am super excited about it. I did not know what was in store for me. When my friends asked me what I was doing this summer I would always respond, "I have an internship offer from IBM". The majority of people would follow with, "Wow, that's so cool! You're so lucky" and it excited me even more. After a busy term at college and even busier short two week trip to my home country, finally, my first day of internship arrived. Everything I had known changed in a very short period of time. New friends, new place and a new summer job! Like many other people, I am uncomfortable with a lot of change coming all at once. However, as a person who always enjoyed new challenges, I decided to embrace this positive change in my life.

Before, I share my first week experience as an Intern, let me introduce International Business Machines abbreviated as IBM and nicknamed "Big Blue" to you all and my role there. The roots of IBM date back to the 1880s, decades before the development of electronic computers. Hard to imagine but IBM's legacy dates back to my great grandfather's period. Primarily, IBM (valued @ $110B) makes and sells software, offers infrastructure services, hosting services, and consulting services in areas ranging from mainframe computers to nanotechnology. Everything that seems impossible is possible for them. Some of the best scientists, engineers and

consultants in the world are IBM'ers. Ginni Rometty is the president, chairman and CEO of IBM. With over 400,000 employees worldwide spread across 175 countries, IBM holds more patents than any other U.S. based technology company and has twelve research laboratories worldwide. It's been ranked as the world's most innovative company for 23 consecutive years. Yahoo! I am a proud IBM 'er today. I cherish this moment.

In its elaborate 136 years of existence, IBM had reinvented itself multiple times and continues to do so. IBM today is leveraging agile practices vastly than ever before. It is rapidly expanding adoption of cloud, analytics, mobile, and social technologies generating a lot of excitement in the industry. IBM management believed that these technologies presented a huge business growth opportunity. Simultaneously, management viewed cloud, analytics, mobile, and social technologies as a disruptive force demanding transformative changes to the way the company worked internally. Management pursued this latest transformation with the belief that the effort would determine business success in the digital economy. As a summer intern with the 'Transformations and Operations' team, I work in the office of the CIO assisting the HRMS team in transforming their legacy HR systems to cloud based application named 'Workday'.

On May 23rd, after a brief welcome message, introduction and completion of paperwork guided by our on boarding specialist team headed by Linda C Carollo, there comes my manager, Bala, welcoming me to the company and walking me to the work space. Having worked in IT for almost five and a half years, my imagination was a typical work space with 'cubicles' all around but to my surprise it's all openwork space. We could clearly sense a great work culture with interesting conversations, laughter and peers pulling each other's legs simultaneously with keyboards and mouse clicking away. Everyone was very excited about us three interns onboarding the team. After a series of introductions and warm welcome messages, we got to collect our brand new 'Mac pros' followed by introductions to a few more peers and senior management teams. It's very inspiring to talk to some of these very accomplished people. Each of them is passionate and honest with a single mission to transform the company. I felt I was exactly where I wanted to see myself.

It did not take me a lot of time to realize that IBM as a global firm and the CIO office with the power to steer the transformation have everything to offer me as per my career interests. Cognitive Computing, Project management, Analytics and a lot of tech savvy work. More importantly, the team is committed to provide these experiences coupled with loads of fun. With few more intro sessions and meetings, we head to the end of week one.

It reminds me of a saying, "Life is change. Growth is optional. Choose wisely." I believe that I made the right career choice by joining IBM. I see an immense potential to grow as a global IBMer and as a 'Blue Agent'. "Positivity and transformation are essential in life, or else you could put yourselves out of existence" is the key message I take away from my week one as an intern.

Project Corner: Predictive Modeling

The BigMart Sales

Business Objective: Big-Mart is a leading retail corporation that operates a chain of supermarkets and grocery stores. In order to optimize overall sales and develop an expansion and go-to market plan in order

to make Big Mart the shopping paradise for buyers, our team developed a predictive model that can assist the management in understanding their overall sales and trends in customer buying behaviors. The aim of the project is to understand the sales data by, test a few hypotheses and finally build a predictive model in order to find out the sales of each product at a particular store, with the various product and store attributes given.

Dataset: The data was the 2013 sales data corresponding to 1559 products across 10 stores of the company in various cities and contains various attributes of each of the products and the corresponding outlet. The attributes of the products include weight of the product, fat content which contains low fat and regular, item type which is the category of the product, maximum retail price and visibility which means the percent of the total area displaying the products to customers. The attributes of the products include establishment year, size of the outlet, outlet type which means the outlet is either a grocery or a supermarket and the outlet location type.

Analysis: After the data pre-processing, the team performed a few test of hypotheses using the independent sample t-test and one-way ANOVA. Post the hypotheses testing, the team ran 6 different models on the dataset with a log transformed dependent variable (Sales) and the top influencing independent variables – Item MRP, Outlet Age, Outlet Size, Outlet Location Type and Outlet Type. The results in the table above are summarized for each of the train and validation datasets

Dataset	Model	R Squared	Error %
Training	Neural Network	74.12%	40.31%
	Decision Tree	74.50%	40.17%
	Random Forest	78.30%	36.98%
	Multiple Linear Regression	71.59%	42.80%
	Lasso Regression Model	71.00%	42.80%
	Ridge Regression Model	71.50%	42.80%
Validation	Neural Network	74.68%	41.03%
	Decision Tree	73.60%	41.87%
	Random Forest	72.70%	43.09%
	Multiple Linear Regression	72.17%	43.54%
	Lasso Regression Model	72.20%	43.54%
	Ridge Regression Model	72.20%	43.54%

Business Insights: From the results of the predictive models developed, we derived the following insights for the management team of BigMart.

1. Increasing the MRP of a product or having products with higher MRP, doesn't necessarily translate into a huge increase in sales. This can be taken into consideration while creating marketing plans and for procuring high cost products for the stores.
2. The management should focus their energies on the medium sized stores followed by small sized stores and try to optimize their sales vs. spending a lot of infrastructure and capital on large sized stores
3. Stores in Tier-2 and Tier-3 cities fare better than the ones in Tier-1. Further, Tier-2 city stores fare better than Tier-3 cities. As such, when deciding to open new outlets, priority should be given to Tier-2 cities. This could also mean that the management can avoid additional overhead costs of expensive land, maintenance in the Tier-1 cities
4. As a store grows older, the sales tend to grow higher. This could be associated with people's increased goodwill and trust in a well-established store compared to a lot of new stores. As such, if the

management is in a dilemma whether to open newer stores or to spend on renovating the old stores, the preference should be given to renovating older stores which are more established.
5. Supermarkets always tend to sell more than the grocery stores. Infrastructural investments, if possible, should be made in supermarket stores vs. the grocery stores. It is statistically better option to open a new supermarket store vs. a grocery store, if one is looking at long term benefits, notwithstanding the initial capital investment in opening a supermarket store.

Team Members: Mir Akram Ali Yadullahi, Mihir Sanghvi, Krishna Chaitanya, Siddharth Kajampadi, Yunqing Zhang

March Madness Mania: Predicting the Winners of the NCAA Men's Basketball Tournament

Business Objective: The goal of our project was to predict the winner of the 2016 March Madness tournament using previous performance data and tournament results.
"March Madness" is the name of the time period in March when the men's NCAA basketball championship tournament takes place. Outside of "March Madness" basketball teams compete during the year in what is known as the "regular season." At the end of the regular season, the sixty-eight (68) best teams are chosen to compete in the March Madness tournament. Each of the 68 teams is provided a seed number based on overall win record, strength of division and team consistency. This seeding selection will determine the line-up of which teams will play in a head-to-head single elimination match.

Dataset: The dataset for this project was pulled from Kaggle - an online data science competition platform. There were several files contained in the dataset for this competition including detailed results from each game played in regular season since 2003, and from every tournament since 1985, as well as team seeds each year since 1985.

Analysis and Insights: In order to understand the datasets, we first performed research on the game of basketball. Basketball is a head to head competition consisting of two teams and one winner per game. Points are earned by throwing the ball into your team's designated goal basket. Balls successfully "dunked" in the basket may earn 2 or 3 points depending on the distance from which the ball was thrown. Teams offensively attempt to score as many points as possible. They also defensively attempt to block the opposing team from scoring. The team with the higher score at the end of gameplay wins.
The data provided by Kaggle included regular season as well as past "March Madness" tournament data for each game played. Data included winning score, losing score, number of overtime periods, number of "three-pointers" per team, number of assists per team, number of offensive rebounds, number of defensive rebounds, etc.

Our group first developed some hypotheses about what might help to predict whether a team will win a game:
(i) Historical % wins of Team 1 against Team 2 will predict probability of future win of Team 1 against Team 2, if sample size (number of games played between Team 1 and Team 2) is great (arbitrarily, n>5).
(ii) Seed difference between Team 1 and Team 2 can predict which team will win.
(iii) The difference in mean point differential of Team 1 and Team 2 can predict which team will win. ("Mean point differential" for Team 1 is calculated by averaging [Team 1 Score - Team 2 Score] for all games played by Team 1)

The first hypothesis was dismissed due to a lack of data. Since teams usually only play other teams within their region or division during the regular season, many teams which made it to the tournament had never played against other teams in the tournament,

so there was no historical precedent that could indicate the probability of future wins. Sample size for team combinations where there was data was too small to be predictive.

This left the second two hypotheses. These were tested by using common models for categorical outcome variables: logistic regression, decision tree and neural net models. For hypothesis (ii), we computed seed difference [Team 1 Seed - Team 2 Seed] and mapped this to either 1 (signifying a win for Team 1) or 0 (signifying a loss for Team 1).

For hypothesis (iii), we computed mean point differential difference [Team 1 MPD - Team 2 MPD] and mapped this to either 1 (Team 1 win) or 0 (Team 1 loss). After testing each hypothesis using different models in JMP, we concluded that the best results were given using seed difference as our only predictor variable in a decision tree model. Error and accuracy results of Decision Tree model vs Logistic Regression model on Test Dataset using seed difference as predictor are shown in the table below.

	Misclassification	Sensitivity	Specificity	Accuracy
Model	Testing	Testing	Testing	Testing
Logistic Regression	35	35	65	65
Decision Tree	32.5	35	81.67	67.5

Comparison of various Model Performances

Using seed difference as the only predictor may seem simplistic, but does make sense. Seed is, after all, a measure of a team's strength and itself is computed using variables which predict how a team ranks amongst other teams. Our model accurately predicted all but 1 of the 16 teams to reach the quarterfinals. It also accurately predicted the top 8 teams.

Team Members: John Birchenough, Taneja Young, Adarsh Anand, Shraddha Sharma, Sishi Yang

Twitter Analytics

Business Objective: Twitter and corporations want to find out what factors increase influence in the social network. Is it the number of followers, the number of posts, retweets or mentions a user receives and sends? Or maybe some other factor? Using this knowledge, companies may be able to adjust their Twitter communication strategy to be more effective. So far, organizations mainly focus on gaining more followers. Any finding that shifts this focus would be valuable.

Data Set: Level of influence in this case is defined as a rather subjective measure. Some people are influenced more by popstars, others by politicians. There is no good way to calculate influence without the involvement of human judgement. This is why Twitter provided a data set, in which each row contains data of two individual users on Twitter (machine generated) and includes a human response which one is more influential. The team used this data to create various models to predict the human judgement (is A or B more influential?) and to identify the most influential factors, by performing stepwise analyses.

Analysis: The team moved on by creating and tuning five types of predictive models: decision trees, discriminant analysis, neural nets, regression, and ensemble models. For decision trees, after looking at the R^2 values, the team decided that four splits were enough (no significant rise anymore). To back it up, it also created the model and compared the accuracy of the model of having only four splits with the accuracy of having a very high number of splits. With a lot more complexity, the 3% more complexity could be added. The team decided that this was not worth adding complexity, especially because the model performs so different on every data set provided. It would have been very realistic that by adding complexity, the model would be overfitted according to the training data. In a similar

fashion, the team dealt with the discriminant and regression analysis, including only few variables to predict vs. all of them. Four variables turned out to deliver a reasonable level of accuracy while still keeping the model simple enough. For neural nets, there was no way to influence which variables were taken into account, but as the table indicates, it clearly is the worst model in terms of accuracy and therefore does not have be considered anyway. Finally, an ensemble model was created using the outputs of all other models and decision trees. However, to achieve a high level of accuracy, many splits have to be made, making it uninteresting because of the high complexity.

Training Data	Decision Tree	Neural Net	Discriminant	Reg
0-0 Correct	2003	2043	1569	1
1-1 Correct	1953	1727	2264	2
Total Correct	3956	3770	3833	3
Percentage Correct	72%	69%	70%	7

Test Data	Decision Tree	Neural Net	Discriminant	Reg
0-0 Correct	2597	2454	1983	1
1-1 Correct	2102	2112	2598	2
Total Correct	4699	4566	4581	4
Percentage Correct	79%	77%	77%	8

Total Right	8655	8336	8414	8
Pct Right Avg	75%	73%	73%	7

Business Insights: The best model to predict human judgement is regression analysis, followed by decision trees and ensemble models. The most important factors on which models base their predictions are follower count, mentions/retweets received, network feature 1 and listed count. This result is valuable for business as it suggests that companies should not only focus on follower count but also keep creating interesting content that creates retweets and mentions which would make their profile more influential.

Team Members: Julian Vrohlings, Mrinal Gaur, Shirley Tarabochia, Phanindra Musunuri, Satinder Sondhi.

Faculty Spotlight

Dr. Moustapha Diaby

Dr. Moustapha Diaby, can you briefly introduce yourself and shed some light on your research areas?

My research in is the areas of Operations Research and Operations Management. In my Masters thesis I developed a Simulation model for routing buses and for deciding the fleet mix and capacity needed in moving students between the multiple campuses of my alma mater, University at Buffalo – The State University of New York. This Simulation model was implemented, and significant yearly dollar savings were reported by the office which had called for the study.

My Doctoral dissertation arose out of a production planning problem which had been brought to two of my thesis advisors (Drs. Stan Zionts and Harish C. Bahl) by a General Motors Corporation supplier who was then located in the Buffalo, NY area. This lead to what, I believe, are still state-of-the-art developments in terms of practical applicability in the area of "Capacitated Lot-Sizing."

Over the past 10 years, a major focus of mine has been the theoretical foundations of scheduling, sequencing, and routing problems. A refereed book I co-authored with Dr. Mark H. Karwan (who had been my third doctoral thesis advisor) on this topic was published in February 2016. This book (entitled "Advances in Combinatorial Optimization: Linear Programming Formulations of the Traveling Salesman and Other Hard Combinatorial Optimization Problems") is about a generalization of previous developments of mine which provide an affirmative resolution to the famous "P versus NP" question, a problem which was classified in 2000 as one of the seven (7) most important

problems for Contemporary Mathematics (i.e., "The Millennium Problems"). One significant aspect of the newly-published book is that it includes fully-detailed, rigorous proofs of the non-applicability of the "impossibility" hypotheses and claims in the pre-existing literature regarding what we are actually able to demonstrate through the computer code implementations. Another major focus of my research in recent years has been the modeling of Loyalty Reward Programs, which are now-a-days very common in almost any kind of industry.

You have degrees in Chemical Engineering, Industrial Engineering and a PhD in Management science/Operations Research. What inspired you to become a professor and what are your expectations from students?

Mathematics has been a favorite subject of mine since a young age. During my senior year in the Chemical Engineering program, I was somewhat dis-satisfied with what I perceived to be a "disconnect" between the theory and practice of the field. I took first courses in Computer Programming and Optimization respectively, and got "hooked" on both immediately ("Love at First Contact!"). I determined that the best possible way for me to combine these two aspects was to pursue an "Operations Research" degree, and the best place for that at the time was the Industrial Engineering (IE) department. I was at the "All-but-dissertation" stage in the doctoral program in IE, when I was confronted with a major decision due to financial issues. My "main" advisor (Dr. Karwan) and the IE Department (which was small at the time) were having financial problems. On the other hand, my "co-advisor" (Dr. Zionts) could offer a (rather generous!) fellowship, but only if I moved to the Management Science/Operations Research program within the School of Business. Part of my reason for accepting this fellowship was the potential advantage I saw then in being able to re-examine some of the same problems as I had in IE, but from a more "real-world"/business perspective. I am happy to have made the decision, because my experience afterwards has proved to me time and time again, that it allowed me to gain a somewhat more "rounded view" on the field of Operations Management (which became the subject area of my Doctoral Dissertation) than the typically-trained Industrial Engineer or the typically-trained Business School person.

Becoming a professor was a natural thing for me, because I enjoy the pursuit of scientific discovery immensely, as well as the challenge of being able to communicate complex ideas in simple, succinct, and clear terms.

Analytics has been around for quite a while now, with many old school algorithms, techniques still the basis for it. Why do you think analytics is so much more relevant in this era?

I think the increasing Globalization, combined with the increasing scarcity of resources, combined with the increasing demands for flexibility and customization, combined with the exponentially-increasing rate of growth of Technology, combined the advent of "Big Data" have created a condition of "Perfect Storm" which is ushering in the time which will eventually come to be known as "The Era of Analytics-Centered Business Decision-Making."

Please tell us about your hobbies and your favorite pastime. How do you prefer spending your weekends, if you are not working, of course ☺?

I enjoy reading, watching sports, and spending time with friends and family.

Alumni Spotlight

Bharath Shivaram

Can you walk us through your professional

journey after graduating from the MSBAPM Program back in 2014?

Soon after graduation, I joined SCIO Health Analytics as Data Analyst, and moved to Deloitte Consulting as Business Analyst in their Innovation practice and I currently work for EMC Corporation as Technology Consultant focusing on Big Data and Cloud technologies.

According to you, how can analytics are leveraged for social entrepreneurship? Can you shed some light on your 'Progressive Insights' initiative?

Progressive Insights, is a Non-Profit organization with a vision to "Drive social and economic progress using public data and analytics". We are passionate about using data and analytics to drive social and economic progress in the developing world. We act as information catalysts in the society by helping not for profit organizations, governments, academic institutions and other think tanks with research and analytics support. At the heart of our organization is the social impact lab which has advanced research and analytics supported by big data technology to store and process large volumes of public data. We are advocates of data literacy in the developing world and consult with governments, non-profits, academic institutions and other think tanks to educate them on the benefits of collecting quality data for effective decision making.

Can you share some important tips/best practices for the current MSBAPM students looking for full-time opportunities, given your understanding of the analytics industry and its demands?

My strategy is to cast a wide net to catch a lot of fishes and choose the best fish to cook. I would suggest the current students to look beyond the data scientist hype and see what meets their prior background experience and their current skills. There are jobs under the umbrella of Information Management which are "process oriented" which have good career scope.

What, according to you, are the top 3 skills essential for a successful career in analytics and project management?

Analytics and project management are two different worlds. Just like men are from mars and women are from Venus. They rarely coincide. Those who like project management are big picture thinkers; they like to see the forest as a whole. Analytics professional like attention to detail and they identify individual trees with a forest and often have difficulty seeing the whole forest. I would suggest the students to take a "complimentary approach" than a "Best fit approach" and focus on skills that are complimentary to their current knowledge and personality.

What do you do in your free time, when you are not being a Technology Consultant!

I think about marriage, I am Kidding! I like to read, play golf, badminton and talk to my parents. I sometimes also enjoy random drives to rural mass and think about love, life and work. https://www.linkedin.com/in/bharathshivaram

Student Spotlight

Deepak Kumar Sisodia

Describe your 10-year professional journey at Adobe Systems

My 10-year long journey at Adobe was very fulfilling and rewarding. I started as an entry level professional and worked on all aspects of a software product -- from testing a product feature to developing automation tools and frameworks to managing an independent team to working with product management in defining and evaluating features. I worked in both "Waterfall" and "Agile" environments. In later years of my career, recruiting and mentoring

new hires became an important part of my job. In 10 years I transitioned from an individual contributor to a decision maker role. Two skills which my work experience at Adobe helped me strengthen are decision-making and problem-solving.

With more than 12 years of work experience, tell us about your motivation to join back school and analytics in general and UConn in specific.

Throughout my career, I was never away from studies. From 2010-2012, I completed my part-time MBA for working managers from IIM Lucknow. It was a very demanding 2.5 years' program. I gained exposure to many new concepts, ideas and paradigms which served me well as my responsibilities at workplace expanded. While doing this program I got introduced to the field of analytics. I did some academic projects also in areas like marketing research. It was during this program that my interest in the field of analytics really piqued. I wanted to do my MBA specialization in analytics but, unfortunately, the option was not available. To build expertise in this field I decided to pursue an MS in Analytics. I chose UConn because it was the only analytics course I found which placed equal emphasis on building business acumen and on pure analytics. Most other schools' curricula focused primarily on data science

Can you describe some specific project management experiences/challenges you faced in your professional career?

Two major challenges which I faced at my workplace were dealing with resistance to change and evaluating team performance. After following the Waterfall model for many years, Adobe adopted Agile in 2012. There was initially a lot of resistance to this change – especially from managers. Being in a Scrum Master role I had several conflicts with the engineering manager. This initially resulted in reduced communication between us, but it quickly became clear to me that in order to effectively resolve disagreements and gain "buy-in", there needed to be more – not less – communication. Engaging in communication on a one-on-one basis really helped me to influence my colleagues to embrace change.

Team performance management is very challenging in practice. Knowledge of recency error, halo effect etc. does help managers to remain unbiased but it is impossible to quantify all aspects of performance. Striking a fine balance between subjectivity and objectivity is critical.

Which industry would you like to join after graduation and why?

I want to join a consulting firm as a Management consultant. I would like to build upon my problem-solving ability so that I can drive decision making at a more strategic level. Being a consultant you have more freedom to speak your mind. You can tell company CEOs what they are doing wrong and in fact this is what you are being paid for!

How do you/did you maintain work-life balance?

Work-life balance varies with time. Your work-life balance will change when you are single vs when you are married. So there is no one-size-fits-all solution. I feel that time management is the key here; however, this is easier said than done. A few things which helped me were – taking time off from the office, having family dinners together, not "taking work home" as far as possible, and regularly getting seven to eight hours of sleep.

Talent of the Month: Subtle Shades of Entrepreneurship – Ashish Gupta

I started my first venture at the age of 22 after graduating as a Mechanical Engineer from Delhi College of Engineering, India. Eventually, I got involved in another start-up during the same year. Both of them were bootstrapped. To tread upon this path, I feel, was a natural inclination as per my demeanor, capabilities, and circumstances. It was a very healthy experience for me and I believe there are many more similar experiences waiting for me ahead. From this roller-coaster ride for over two years, I've learned quite a few things that I'd like to share with the readers. A start-up could be at any scale and there are many articles that talk about the financial challenges, market challenges and technical challenges in all kinds of start-ups. However, only a few tell us how to make personal level trade-offs, which are the common challenges faced by the people involved in this job. The best way you can learn that is by your own experience. Only after starting two start-ups and getting involved in a couple of more, I've realized why the trade-off at personal level is equally important. In a start-up, there isn't any sooner stopping once the commitment is made. The essence of success is very abstract in earlier stages. So abstract that you might miss apprehending a few nuggets, really! There isn't a paycheck coming every month that suffices you or makeable developments following one after the other that could motivate one to keep going. In most cases, you'd be nurturing what you have created – trying to capture/create the market or dealing with some technical issues. Sometimes, people aren't completely ready to make this commitment and the interesting part is that they themselves don't know about it. To survive this period, you have to have a very strong motivation, a true dedication and an honest commitment. And what really helps achieving this is a very clear mindset about your long-term goals, your expectations and your sacrifices. A personal level trade-off – what you expect from your life, when and how, do matter in this painfully beautiful journey! As it goes by, the last thing you want your company to suffer from is "you". But, just wanting it the least doesn't rule out this possibility, Right? It's not that what you achieve defines you but what you do! And you can really achieve something from anything when you do it with all your heart. Maybe this aspect was more conspicuous to me because of my unnatural timing for pursuing this endeavor but I tell you, be mindful of it.

The success during the early stages of your start-up could be the patience (which you already have or you develop) to wait for your upcoming successes. Wise and honest personal level trade-offs definitely help the cause.

"Summer is Here" …

Protect yourself!

Who doesn't want to be out this summer to the beach or barbequing? While summer is fun, make sure you don't have an overexposure to sun like sunburns, premature aging and skin cancer.

Some tips to keep yourself safe this summer:

Avoid hours between 10 a.m. to 4 p.m. outside. They are most hazardous UV exposure in USA.

Apply Sunscreen: Use sunscreen with sun protective factor (SPF) 15 or higher, and both UVA and UVB protection. Don't forget to check the expiry of your lotion.

Clothing: Wear long-sleeved dry shirts to protect yourself. Try get some T-shirt with SPF rating lower than 15. Swim leggings are a great option for those who like running on the beach; you can get them wet

and protect yourself from strong sun rays at the same time.

Hats: Wear a hat that has a wide brim to cover your face, ears and neck. A tightly woven canvas fabric hat is the best. A darker hat adds more protection.

Sunglasses: It's a must as it protects from cataracts. Sunglasses that block both UVA and UVB rays offer the best protection. Most sunglasses sold in the United States, regardless of cost, meet this standard.

Look for sunglasses with polarized lenses to protect your eyes; this fun two-tone style looks great with everything from a white beach dress to a printed top and jeans.

Seek Shade: Even if outings, seek shade in midday hours. Reduce damage by seeking shade from an umbrella or tree.

Rash guards: They are an on-trend item, so you can feel confident sporting one wherever you go this summer. Classic surf brands, such as Roxy, have great sporty prints and color blocked styles. Also look for items labeled "swim shirt" or "sun shirt" to widen your search.

Scarf: A chic addition to any beach-inspired outfit, this scarf also will protect your neck, an area often overlooked when it comes to sun protection. Lightly spritz the scarf with cold water for a nice cooling effect when the temperature rises

Have a safe summer!

"Eat Fresh, Eat Healthy!"- The Salads

Thai Cucumber Salad

"This sweet and tangy summer salad of cucumber, cilantro, and peanuts with just a hint of heat is always a hit at picnics and potlucks since it doesn't need to be refrigerated and you're pretty much guaranteed to be the only one bringing this dish!"

Ingredients

3 large cucumbers, peeled, halved lengthwise, seeded, and cut into 1/4-inch slices

1 tablespoon salt

1/2 cup white sugar C&H Sugar Pure

1/2 cup rice wine vinegar

2 jalapeno peppers, seeded and chopped

1/4 cup chopped cilantro

1/2 cup chopped peanuts

Directions

Toss the cucumbers with the salt in a colander, and leave in the sink to drain for 30 minutes. Rinse with cold water, then drain and pat dry with paper towels.

Whisk together the sugar and vinegar in a mixing bowl until the sugar has dissolved. Add the cucumbers, jalapeno peppers, and cilantro; toss to combine. Sprinkle chopped peanuts on top before serving.

Summer Corn Salad

"This fresh and flavorful salad features buttery yellow corn tossed with chunks of tomato and onion with a fresh basil vinaigrette."

Ingredients

6 ears corn, husked and cleaned, white and Bi-Color Corn

3 large tomatoes, diced

1 large onion

Diced 1/4 cup chopped fresh basil

1/4 cup olive oil

2 tablespoons white vinegar

Salt and pepper to taste

Directions

Bring a large pot of lightly salted water to a boil. Cook corn in boiling water for 7 to 10 minutes, or until desired tenderness. Drain, cool, and cut kernels off the cob with a sharp knife.

In a large bowl, toss together the corn, tomatoes, onion, basil, oil, vinegar, salt and pepper. Chill until serving.

Events and Workshops

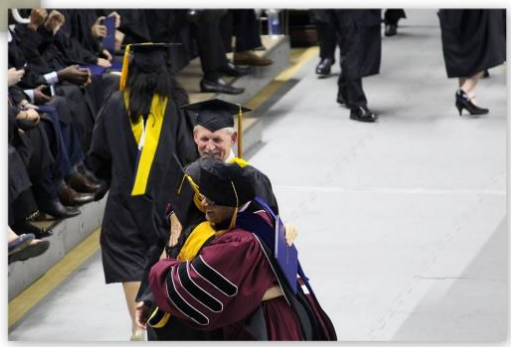
May MSBAPM Graduation Ceremony

Come May and comes the season of graduation! A happy day to see 55 students from the MSBAPM program, graduate and a sad day to know that we bid adieu to our friends, classmates and team mates. May 2nd marked one of the brightest event in the MSBAPM calendar – the MSBAPM graduation ceremony, attended by students in great numbers, faculty and the program management team of MSBAPM. The event was marked by motivational speeches from our esteemed faculty team, followed by the graduation ceremony, calling out to all the 55 students, who will now cap the proud logo of a UConn alumni. The event was also marked by conferring merit scholarships to 15 students in the program, as listed in the table below. As they say, pictures say more than a 1000 words!

May 2016 - Student Merit Scholarship Recipients		
Bhat,Akhil	Joshi,Chinmay	Nie,Wandi
Cao,Chunxue	Kamath,Vikram Markad	Parikh,Harshal
Feng,Shiwan	Kanwar,Shikha	Raju,Abinaya
Ghorpade,Akhil	Khemka,Aayush	Yang,Jingwen
Jain,Rahul	Lemp,Walter Christian	Zhang,Mozhi

UConn Commencement 2016

Commencement at the University of Connecticut is a time of ceremony and pageantry, a time for marking academic achievements, and a time to begin the next step in life. Family and friends gather to witness the formal end to a student's college years; faculty, staff and trustees join in the tradition of conferring degrees; and students participate in an annual rite that marks their individual accomplishments. It was the weekend of May 7 and 8, 2016 – a bright and sunny weekend, marking the beginning of a capstone experience and defining the start to the students' new role as alumni – valued members of Husky Nation. More than 8,700 students across all the UConn campuses graduated from the University this spring. The ceremony was marked by a great commencement speech from the exceptionally talented and innovative film director, screenwriter, producer, army veteran - Oliver Stone. 55 students from the MSBAPM program graduated in the spring-2016. The Gampel pavilion in Storrs was lit with hundreds of students, family, faculty, the management and the beaming joy on the faces of the graduating class, who thoroughly deserved a sense of UConn pride after their exceptional journey of learning and fun at the MSBAPM program. As they move out into the 'world' and wear the 'alumni' cap of UConn, we wish them all the very best in their endeavors. As said by Oliver Stone in the commencement speech - ***'Listen to the wind. The answer might blow right past you'!***



Newsletter Editors:

Monika Katariya

Alekhya Reddy

Mir Akram Ali Yadullahi