Introduction to Business Analytics
Presenter’s background

- Detection and estimation
  - Quantization of prior probabilities for hypothesis testing
  - MCMC inference
  - Regression
  - Kalman filtering with matrix factorization*

- Machine learning
  - Margin-based classification methods
  - Dimensionality reduction
  - Ensemble classification methods
  - Application to sensor networks

- Sparse signal representation
  - Structured overcomplete dictionaries
  - Application to synthetic aperture radar

- Medical image segmentation
Presenter’s business background

- A one week ‘microMBA’ class
- Two years experience working on business applications at IBM Research

- I didn’t know any of this stuff two years ago and continue to learn every day
- Skewed coverage of business analytics – more of an IBM perspective
  – Especially projects I have worked on
Participant’s background?

Source: http://www.davidairey.com/images/typography/who-are-you.jpg
Agenda

What is business analytics

Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems
What next?
Some businesses operate in the same way that they always have.

http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/8114961/Alliance-Boots-plans-major-push-into-China.html
Some businesses operate in the same way that they always have

1810s

2010s

Source: http://milkmiracle.net/2010/12/18/acland-in-hindoostan
Some businesses operate in the same way that they always have
Some businesses now collect and examine some data
Some businesses are integrating data-driven algorithmic recommendations into their decision-making processes

- These organizations are adopting *business analytics*

- Business analytics is all about the decisions that go into running a business
  – Not quantitative finance or investing in stocks
  – Things like:
    - Hiring, retention, attrition
    - Buying, selling, marketing
    - Staff deployment
    - Strategy
What is business analytics

- Encompasses methodologies from applied mathematics, applied probability, applied statistics, computer science, and signal processing for using data to gain insight into business performance and drive business planning

- Solutions primarily used as decision support systems or as components of decision support systems to aid salespeople, executives, and other organizational leaders in business decision-making tasks

- Not the same as operations research
  - Operations research: technical area of study focused on constrained optimization
  - Management science: application of operations research to business
  - Business analytics: application of statistical signal processing, etc. to business

Business decisions

- How likely is client X to buy product Y?
- Which product should we recommend next?
- What is a “realistic” view of opportunity by client?
- Are there accounts where there is significant untapped revenue?
- Which clients are “at risk” of going to our competitors?
- What kind of salesforce do we need to deliver on our targets?
- Which units are not performing at par?
- Which sellers might miss their quota?
- How can we align sellers with client opportunities to achieve maximum revenue impact?
Business decisions (continued)

- Who are the influencers for this product in the marketplace?
- What is the optimal marketing campaign to deploy?
- Should we hire this employee?
- Which employees are at risk of voluntarily leaving the company?
- How many employees do we need to hire now so that we can achieve our production goals six months from now?
- What kind of raises or promotions should we offer to retain our talented employees?
- Will outsourcing help?
- Which company should we acquire to expand our customer base?
Where is the signal processing?

- As we will see in this tutorial, answering such business questions can be approached using appropriate data/signals as raw materials and signal processing techniques to transform the raw data into meaningful descriptions, predictions, recommendations, and actions
  - Not always obvious how
  - Some questions require existing algorithms and techniques, but others require the innovation of new techniques
  - If you start getting involved with business, you start to see signal processing problems everywhere

- Signal processing has been applied to radar, communications, speech, music, images, video, astronomy, geophysics, remote sensing, biology, medicine, etc.
  - Business applications are a great source of interesting problems and really need the injection of signal processing expertise
  - New frontier for signal processing
Agenda

What is business analytics

Signals and systems view of a business

A detailed example: sales team job role mix

Perspectives on business analytics

Dose-response problems

Binary classification problems

Relational and graph-based problems

Quantile problems

What next?
Business

- Before getting to the analytics, let us consider what a business is and does, and why
  - Take shareholder’s investment and bank loans
  - Use that money to make or buy something
  - And also sell and support it
  - From selling, earn revenue
  - Use to repay debt and to profit
Business as a system with feedback (simplified)

- The signal that flows in this system is money or funds

The signal that flows in this system is money or funds.
Definitions

- **Revenue**
  - The funds received by the business from customers for products or services

- **Cost of goods sold**
  - The cost of specifically making the products or providing the service
    - Also the wholesale cost of products sold by a retailer

- **Selling, general, and administrative expenses (SG&A)**
  - The expenses associated with salespeople, marketing, advertising, research and development, administration, etc.
  - All expenses other than cost of goods sold
Long-term goals

- High return on assets
  \[ \text{ROA} = \frac{\text{profit}}{\text{assets}} \]
  - Effective use of everyone’s money

- High return on equity
  \[ \text{ROE} = \frac{\text{profit}}{\text{equity}} \]
  - Effective use of the owners’ money

- Simultaneously desire:
  - High revenue
  - Low costs
  - Low expenses

Short-term goals

- Meeting earnings per share targets
- Meeting revenue targets
- Meeting cash flow targets

Role of business analytics

- Goal of business decision making is to achieve high revenue, low costs, and low expenses
- Business analytics supports this decision making in various ways
- A couple of examples
  - If the business would like to keep selling expenses the same but increase revenue, business analytics can be used to recommend changes in the deployment of the salesforce so that sales teams are better matched to customers and can thus sell more to customers (higher revenue)
  - If workers of the company are voluntarily leaving for higher paying jobs and there are significant salary premiums and onboarding costs associated with hiring replacement workers, then analytics-recommended targeted raises can retain existing workers, thereby avoiding the high costs associated with hiring
Agenda

What is business analytics
Signals and systems view of a business

A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems
What next?
Sales team job role mix

- Utilizing expenses in a good way to increase revenue
- The salespeople of large companies approach potential clients in teams
- Different salespeople have different job roles
- Two broad categories of seller job roles
  - Technical sellers provide details about products
  - Non-technical sellers build relationships
- High-level question: What is the most effective ratio of technical to non-technical sellers to maintain within the overall organization?
  - Clients are too heterogeneous
- Low-level, operational question: What is the right team for a particular client opportunity?
  - Data is extremely noisy data
- Middle-level question: What is the best team composition for clusters of similar clients?
  - Maximize transactional revenue
Sales team job role mix data

- Ideally, use historical data on sales team composition and transactional revenue to learn relationship between the two and generalize for future prediction and recommendation.

- In practice, it is very rare for a business to have exactly the perfect data to answer the question of interest.

- The particular business we studied maintained the following data:
  - For each seller
    - Job role
    - List of clients (sales territory)
    - Aspirational revenue for the year from each client on his or her list
  - For each client
    - Attributes such as region, segment, sector, industry, coverage type
    - Actual transactional revenue for the year (not broken down by seller or job role)
Surrogate measure for sales team composition

- If Seller 1 has four clients in his territory with aspirational revenues:
  - Client $i$: $100,000
  - Client $j$: $200,000
  - Client $k$: $50,000
  - Client $l$: $50,000
- Then the assumption is that Seller 1 will spend 1/4 of his time on Client $i$, 1/2 of his time on Client $j$, and 1/8 of his time each on Clients $k$ and $l$
  - Full-time equivalent (FTE)
- Seller 1 (technical), Seller 2 (technical), and Seller 3 (non-technical) all have Client $i$ in their territories
  - FTE of Seller 1 spent on Client $i$ is 1/4
  - FTE of Seller 2 spent on Client $i$ is 1/5
  - FTE of Seller 3 spent on Client $i$ is 1/2
  - Total FTE spent on Client $i$ is 1/4 + 1/5 + 1/2 = 0.95
- Sales team composition technical to non-technical ratio $d_i$ assumed to be ratio of FTEs
  - $d_i = (1/4 + 1/5)/(1/2) = 0.9$
Formulation

- Now, for each client $i = 1, \ldots, n$, we have
- Client attribute vector $x_i \in X$
  - Contains values such as region, segment, sector, industry, coverage type
- Sales team composition $d_i \in [0, \infty)$
  - Surrogate from previous slide
- Transactional revenue $y_i \in [0, \infty)$

- Using data samples $\{(d_1, y_1), (d_2, y_2), \ldots, (d_n, y_n)\}$, estimate functional relationship between transactional revenue and sales team composition: $y(d)$
  - Relationship may be different for different groups of clients based on feature vector $x_i$
  and may want to do separate estimates: $y^{(k)}(d), k = 1, \ldots, m$
Nonparametric regression

- The estimation problem calls for data smoothing or regression of some sort
- Do not want to or need to assume any specific parametric form for relationship
- Nonparametric kernel regression is a good option for this problem
Nadaraya-Watson kernel regression

- Nadaraya-Watson formulation is the standard technique:
  \[ \hat{y}(d) = \frac{\sum_{i=1}^{n} y_i K\left(\frac{d - d_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{d - d_i}{h}\right)} \]
  where \( K(\cdot) \) is a kernel function such as the Gaussian kernel or the Epanechnikov kernel and \( h \) is the bandwidth

- Bandwidth selection by standard techniques such as plug-in rules of thumb

- For separate estimates of groups according to attribute vector, such as \( \{ i \mid x_i \in A \subseteq X \} \):
  \[ \hat{y}^{(k)}(d) = \frac{\sum_{i \in A} y_i K\left(\frac{d - d_i}{h}\right)}{\sum_{i \in A} K\left(\frac{d - d_i}{h}\right)} \]
Some results on real data from an actual business

- Will show results on real-world data from the sales organization of an actual business from some recent year
- Certain details about the data cannot be shared due to its proprietary nature
  - Slightly obfuscated

- Number of clients $n$ is approximately 3000
- Total number of salespeople in the organization is approximately 500
All clients (no separate estimates according to attribute vector $x_i$)

Transaction revenue maximized at ratio 1.32.
Adjusted transactional revenue maximized at 1.24.
Teams should spend 25% or 30% more technical effort than non-technical effort.
This is overall view, dominated by certain types of accounts. Interesting differences appear when analysis done for different client-types separately.
Core clients

Transaction revenue maximized at ratio 1.13.
Adjusted transactional revenue maximized at 1.18.
Not too different from view on all accounts.
Slightly less technical effort required with core clients than seen with all clients.
Clients with whom sales could grow

Transaction revenue maximized at ratio 0.36 (3 non-technical sellers and 1 technical seller on a team), but also has a secondary maximum at 1.33.

Adjusted transactional revenue swaps the importance of the two maxima. “More bang for your buck” at 1.47, but maximum also at 0.32.

With high technical effort, less total effort is needed. However, very high non-technical effort teams are also successful.
Clients whose business is taken opportunistically

Two different types of teams are successful at these clients: those with very high technical composition (~2.3 ratio) and those with low technical composition (~0.7 ratio).
Almost even technical to non-technical sales teams, perhaps slightly more non-technical, are successful in the public sector.
Industrial sector clients

More technical effort is needed to successfully sell in the industrial sector.
Teams with higher technical composition should be deployed in the communication sector to see whether a higher ratio is even better.
Computer services sector clients

There are two groups of clients in the computer services sector: those that respond to a more non-technical sales team and those that respond to an extremely highly technical team.
In the distribution sector, both high non-technical teams and slightly more technical teams bring in a good amount of revenue.
More technical teams are more successful in the financial services sector, but a group of clients responds to more non-technical teams as well.
Sales team job role mix summary

- We see similar but different team composition-revenue relationships for different types of clients
  - Give insights for sales staffing
- The leaders within this sales organization expected to see different results before we performed the analysis
- At the very least, this organization is reevaluating their assumptions with regards to staffing

- Hopefully this example whets your appetite to see more
- There are shortcomings of the question asked and the approach taken
  - Will get to them later on in the tutorial
- Sets us up to discuss a general class of problems that arise in business analytics: dose-response problems
Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix

Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems
What next?
History from a managerial perspective

- Industrial efficiency and scientific management, early 1900s
- Decision support systems, 1960s
- Data warehousing, late 1980s
- Business intelligence software, 1990s–2000s
- Descriptive, predictive, and prescriptive business analytics, 2000s–
Business intelligence

- Enterprise resource planning (ERP) systems
  - Finance and accounting
  - Human resources
  - Manufacturing
  - Supply chain management
  - Customer relationship management (CRM)

- Systems that present trusted *historical and current* information in a clear way

- Database systems

- Reporting tools
  - Summarization and visualization of historical data
  - Ability to ‘slice and dice’ and ‘drill down’
  - Dashboards
Business intelligence dashboards

- Graphical representations of the underlying state of the business

Source: http://www.projectmanager.com/project-dashboard.php
Business analytics moves beyond business intelligence

- Not just reporting of historical and current raw data
- Statistical signal processing, machine learning, and data mining
  - Descriptive analytics
    - Transforming data and signals into more meaningful forms
    - Metrics
  - Predictive analytics
    - Estimating future events generalizing from historical data
    - Insights
  - Prescriptive analytics
    - Optimizing actionable inputs
    - Recommended actions and decisions
Besides good statistical signal processing

- How many times have you heard the story of the analyst who comes up with a great concept or model, but the affected decision makers have never heard anything about it, cannot understand it or are unable to ‘operationalize’ it?

1. Understand how the analytics will be used and develop solution accordingly
   - How the decision-making process will need to be changed to take analytics solution into account

2. Transparency, trust, confidence
   - Solution should explain why a recommended action is appropriate in terms that decision makers and stakeholders can understand

3. Results should include a clear interpretation of what action is recommended
   - Prevent information overload

For nearly two decades, Watson had been selling IBM’s tabulating machines as devices that could count and sort tangible quantities – money, inventory, troops and so on. As Wood saw it, that gave IBM a big potential market, but ultimately a limited one.

Wood switched gears and began explaining how IBM machines could be used to measure intellect and psychology. Anything could be represented by mathematics, numbers and formulas. Biology, astronomy, physics or any other science could be aided by IBM machines. Never had Watson considered that numbers could be used to represent and simulate absolutely anything.

Now, we are applying signal processing methods developed for astronomy, communications, physics, and other sciences back in the business realm, and developing new algorithms for the new problems that are encountered.

Humans and their relationships

- One of the key factors that we have to account for is the human element
  - Does not arise when counting money and inventory
- Uncertainty
- Relationships
  - Buying and selling
- Growth and change
  - Learning new skills

Adoption of business analytics actually does help

Note: Respondents were asked about their organization's application of analytics to the activities listed above. A score of 1.0 indicates an equal likelihood of applying either analytics or non-analytic methods, while a score of 0.0 indicates a tendency to use non-analytic methods.

Adoption of business analytics actually does help (continued)

- 6% higher output and productivity than would be expected given other investments and information technology usage

Obstacles to adoption

- If business analytics adoption is so beneficial, then why haven’t all businesses adopted it?

- Most organizations make the mistake of believing that applying analytics is 90 percent math and 10 percent organizational change management with employee behavior alteration. In reality it is the other way around; it is more likely 5 percent math and 95 percent about people.

- So what continues to obstruct the adoption rate of analytics? Major ones are social, behavioral and cultural issues, including people’s resistance to change, fear of knowing the truth (or of someone else knowing it), reluctance to share data or information and a “we don’t do that here” mindset.

Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics

Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems

What next?
The basic idea of dose-response

- **Term dose** is usually used in the context of health and medicine
  - Amount of treatment or drug administered to affect the state of the patient in a desired way
  - The dose of the treatment for a particular patient is determined by a physician, the decision maker, according to a known relationship between the dose and its response on patients
  - Actionable

- **Dose-response function**

- Need not limit to the field of health and medicine

Source: http://www.originusinc.com/Application.htm
Sales team job role mix as a dose-response problem

- In the sales team composition problem, we have three types of variables per client
  - Attributes
  - Dose (team composition)
  - Response (transactional revenue)

- Sales team composition is a dose because it is actionable
  - Choice of the decision maker, i.e. sales manager or executives

- Industry sector of client is an attribute because it is not actionable
  - Sales manager has no action that would change a client’s sector from distribution to financial services

- Transactional revenue is the response because it is the quantity being measured and optimized
  - Sales manager cannot act on it directly, but only indirectly through the dose
  - Response to dose is affected by attributes
    - E.g., different dose-response functions for different industry sectors
Dose-response formulation for business analytics

- Objects $i = 1, \ldots, n$
- Attributes $x_i$
- Dose $d_i$
- Response $y_i$
- Dose-response function $y(d)$

- Many business analytics problems fit into this framework
  - Not necessarily with $d_i, y_i \in [0, \infty)$
- Appropriate formulation because it provides clear actionable insights to decision makers
- Different available data and assumptions lead to a variety of algorithms for estimating $y(d)$ in different problems
Example 2: An organization to complete routine pre-sale activities

- Salespeople spend some of their time performing routine pre-sale tasks such as proposal preparation, solution design, pricing approvals, and contract approvals rather than interacting with customers.
- A business has organized a specialized group to handle these pre-sale tasks for the sellers.
- The business would like to know whether this specialized group is effective and if so, for which sales opportunities.
- Objects: sales opportunities.
- Attributes: revenue size, business complexity, cross-brand or not, client attributes.
- Dose: $d_i \in \{0,1\}$ specialized organization used or not.
- Responses:
  - $y_i \in [0,1]$ probability of completing the sale (win rate).
  - $y_i \in [0,\infty)$ time duration from opening to closing of opportunity (cycle time).
Example 2b: An organization to complete routine pre-sale activities

- Instead of looking at effectiveness of specialized organization per opportunity, look at it per seller
- Objects: sellers
- Attributes: field sales or telesales, business unit
- Dose: \( d_i \in [0,1] \) fraction of opportunities sent to specialized organization
- Responses:
  - \( y_i \in [0,1] \) probability of completing sales (win rate)
  - \( y_i \in [0,\infty) \) number of opportunities carried at a time
Example 3: Business impact of outsourcing

- Outsourcing
  - Any task, operation, job or process that could be performed by employees within an organization, but is instead contracted to a third party for a significant period of time
  - (Not necessarily offshoring)
  - Relief of resource shortages
  - Ability to concentrate on core business
  - Monetary savings
  - Flexibility

- Is outsourcing effective at decreasing SG&A and increasing earnings before taxes (EBT)?

- Objects: companies
- Attributes: firmographics
- Dose: \( d_i \in \{0, 1\} \) initiation of outsourcing engagement
- Responses: \( y_i(t) \)
  - time signal of SG&A
  - time signal of EBT
Example 4: New seller productivity

- Salesforces are in constant flux with people leaving and joining
- Each event affects headcount immediately but effect on sales capacity is not immediate
  - A university new hire may have zero productivity, i.e. sales revenue, for a few months and then gradually ramp up to average sales force productivity over the following few months
- True sales capacity, which is the effective number of sellers, should be used for planning and seller deployment optimization rather than headcount
- Objects: employees (specifically sellers)
- Attributes: whether a new hire from a university, with experience, transfer from a different business unit, or from an acquired company
- Dose: \( d_i \in \{0,1\} \) hiring
- Response: \( y_i(t) \) time signal of sales capacity
Example 5: Product recommendation

- Cross-selling, up-selling, and recommending appropriate products to customers are important ways to increase revenue
- Objects: customer purchases
- Attributes: if business customers, then firmographics; if consumers, then demographics
- Dose: $d_i \in \{1, \ldots, \ell\}$ products available to sell
- Responses:
  - $y_i \in [0,1]$ probability of completing the sale (win rate)
  - $y_i \in [0,\infty)$ expected gross profit (probability $\times$ unit gross profit by selling product $d_i$)
Example 6: Talent retention

- If workers are voluntarily leaving a company for higher paying jobs elsewhere, and there are significant salary premiums and onboarding costs associated with hiring replacement workers, then retaining existing workers avoids the high costs associated with hiring.
  - Also no productivity losses

- Objects: workers

- Attributes: region, business unit, job role, length of service, performance rating

- Dose: \( d_i \in (-\infty, +\infty) \) percentage away from average salary of peers

- Response: \( y_i \in [0, 1] \) probability of voluntarily leaving company
Example 6b: Talent retention

- In addition to the effect of how much underpaid or overpaid a worker is on probability of voluntarily attrition, the mere act of receiving a promotion or an increase in salary transiently affects this probability
- Objects: workers
- Attributes: region, business unit, job role, length of service, performance rating
- Dose: \( d_i \in \{0,1\} \) promotion or salary increase
- Response: \( y_i(t) \) time signal of probability of voluntarily leaving company
Example 7: Post-disaster philanthropic giving

- After major disasters like earthquakes, non-profit relief agencies receive a flood of donations, but the individuals who donate do not always continue to give later on
  - Non-profit organizations are also businesses that need to be managed and are starting to use business analytics

- Agencies solicit donations by various channels continually

- Objects: members of the general public

- Attributes: demographics

- Doses:
  - \( d_i \in [0, \infty) \) magnitude of disaster
  - \( d_i \in \{1, \ldots, f\} \) interaction (phone call, short letter, letter with glossy pamphlet, etc.)

- Response: \( y_i(t) \) time signal of expected donation amount
Catalogued dose-response examples

- Quite a variety of business applications: salesforce analytics, workforce analytics, marketing
- Quite a variety of data types: continuous scalar, binary and \( \ell \)-ary discrete scalar, functional
- Not obvious that all fit into dose-response format
  - Require variety of estimation techniques, including novel ones
  - Different ‘hammers,’ not just one hammer looking for nails

- Agenda for remaining dose-response section
  - Return to sales team job role mix to address shortcomings
    - Different response variable
    - Decision tree for automatic grouping
    - Dynamics
  - Available data and estimation techniques for the other catalogued examples
  - Going from predictive analytics to prescriptive analytics

- After that, discuss problem types other than dose-response
A few shortcomings of job role mix formulation

- The response variable in the first go around was transactional revenue
- A client of a particular size will produce transactional revenue in a particular range regardless of the sales team
- Ratio of actual transactional revenue to aspirational revenue is a better measure of effectiveness than just the transactional revenue
  - Lift given by sending the appropriately composed team
- Even better than transactional revenue in the numerator of the ratio is (transactional revenue – sales expense)
  - More directly related to operating profit
  - Use $y_i = (\text{transactional revenue} - \text{sales expense})/(\text{aspirational revenue})$
  - Sales expense is closed form function of seller FTEs
- Only obtained the expected value of the dose-response function
  - Did not characterize the uncertainty around $\hat{y}(d)$
  - Can easily use slightly modified kernel regression to estimate $\sigma(d)$
    - Standard deviation of response at a particular dose
- Just blindly obtained dose-response functions for different groupings of clients
  - All segments, all sectors, all regions, …
A decision tree to partition attribute space according to dose-response behavior

- Instead of blindly performing all possible cuts of the clients, partition $\mathcal{X}$ in an intelligent manner so that each partition has a common dose-response function that is different from all other partitions.
- Have developed a new decision tree method to split $\mathcal{X}$ in a top-down manner:
  - Split best individual dimensions of $\mathcal{X}$ to create leaves with homogeneous dose-response relationship.
  - Binary tree.

A decision tree to partition attribute space according to dose-response behavior (continued)

- A proposed split along an attribute dimension, e.g. communication sector and distribution sector clients in one branch and clients from all other sectors in the other branch, will create two populations of data points \{ (d_i^A, y_i^A) \} and \{ (d_i^B, y_i^B) \}

- Split criterion
  - Find kernel regression estimate from all data points at node
  - Calculate residual between A data points and estimate, and residual between B data points and estimate
  - Calculate Cramér-von Mises statistic between these two residuals
    - Squared L^2 distance between cumulative sums of residuals

- Calculate split criterion for all possible dimensions and splits, and split on one with largest statistic value
  - If all possible splits give small value of statistic, do not split
    - Leaf of tree

A decision tree to partition attribute space discussion

- By splitting on attributes, the groupings that are found are actionable
  - E.g. decision makers can institute policies that clients in this region, industry, and segment will be served by this team composition
  - Hiring/attrition and salesforce planning can be performed at those levels

- In results on the same data as shown earlier
  - First split is on coverage type: clients with the sales team having desks on site, and others
  - Among others, next split is on segment: core clients and those whose business is taken opportunistically in one branch, and ones whose business could grow on the other

One more shortcoming of job role mix formulation

- Assumes that behavior of a client in the past will carry forward to the future
- Assumption is alright for groups of clients generally, but not when looking at individual clients
  - Trajectories of behaviors
    - A client may need a team with more non-technical sellers at first and a team with more technical sellers later as the relationship develops
- We have developed a novel dynamic model and algorithm for estimation that can be used to overcome this shortcoming in the sales team job role mix problem
  - Matrix factorization
  - State space model
  - Kalman filter and smoother
  - Expectation maximization
- See talk by John Z. Sun, Thursday, March 29, 11:50 am, Room B-2

Specialized organization for pre-sale activities

- For each sales opportunity, the available data includes whether the opportunity was sent to the specialized organization, dates of progression through the sales pipeline (identified, validated, qualified), won/lost, revenue size, business complexity, cross-brand or not, list of sellers involved, information about sellers

- Example 2
  - Straightforward estimation of dose-response $\hat{y}(d=0)$ and $\hat{y}(d=1)$ for different attribute values by simply averaging time duration or won/lost for all opportunities that satisfy attribute value
    - Separate estimation for those sent and those not sent to specialized organization

- Example 2b
  - Rearrange opportunity data so it is organized by seller
  - $d_i$, fraction of the seller’s opportunities sent to the organization, is easily calculated
  - $y_i$, win rate, is simply the average won/lost for all of the seller’s opportunities
  - $y_i$, number of opportunities carried, is figured out by taking a small window in time and counting how many of that seller’s opportunities are in an active stage of the pipeline
  - Then $\hat{y}(d)$ can be estimated using kernel regression, separately for different seller attributes
Business impact of outsourcing

- Publically traded companies are required by law to release their income statements
  - Quarterly (every three months) data on SG&A and EBT but with missing values
- Large strategic outsourcing deals are announced in press releases
  - Dates and values of contract signings
- Thus have time series about outsourcing deals \( \{y_1(t; d_1=1), \ldots, y_n(t; d_1=1)\} \)
  - Time measured in quarters of years; \( t = 0 \) is deal signing date
  - Look at growth rate of SG&A, EBT instead of raw quantity

Business impact of outsourcing (continued)

- If we already have \{y_1(t; d_1=1), \ldots, y_n(t; d_1=1)\}, then what’s the estimation problem?
  - Extremely noisy
  - Want to find similar structure or shape of response signal in different clients
  - Variability across clients in amplitude, time scale, time delay

\[ y_i(t; d_i=1) = A_i \hat{y}(b_i t + c_i) + \text{noise}(t), \quad i = 1, \ldots, n \]
  - Coefficients \([A_i \ b_i \ c_i]^{T}\)
  - Common structure \(\hat{y}(t; d=1)\) polynomial interpolating spline
  - Gaussian noise

- Bayesian priors for \(A, b, c\), same for all \(i\)

- Markov chain Monte Carlo inference of coefficients and common structure
  - Gibbs sampling
  - Gaussian samples for spline parameters
    - Mean and variance determined by spline interpolation formulas
  - Metropolized independent sampling for coefficients

SG&A growth rate results

\[
\hat{y}(t; d=1)
\]

EBT growth rate results

\[ \hat{y}(t; d=1) \]

Business impact of outsourcing discussion

- Outsourcing provides temporary relief for a year or a year and a half
- Of interest to see if there are different response functions for companies with different attributes
- This sort of analysis is performed across companies
  - Not specifically useful for decision makers within one particular company
- Usefulness is to strategic outsourcing vendors
  - If business impact can be shown beneficial, a vendor can better market its services to prospective clients
- Reports presenting these types of results are useful marketing collateral

New seller productivity

- Revenue of individual sellers is not recorded in a good way to enable direct examination of their productivity
  - Mainly due to stacking
    - The entire sales team’s revenue gets recorded for each individual without breaking down their contributions
    - Need an estimation procedure to extract individual behavior of new sellers from aggregate revenue data

- Available data in this problem includes:
  - Quarterly overall transactional revenue for the entire business unit
    - Adjust for seasonal effects (many businesses sell significantly more in the fourth quarter)
    - Interpolate quarterly to monthly
  - Total seller headcount each month
  - Headcount dynamics each month
    - Attrition count
    - New seller count

New seller productivity (continued)

- $r$: total revenue in business unit (adjusted for seasonal effects)
- $\Phi$: convolution matrix with headcount dynamic entries
  - 4 different classes of additions, 1 class of attrition, ‘old-timers’
- $\hat{y}(d=1)$: sales productivity profiles for the different classes
- Consider $l_\infty$ norm (maximum error) for data fidelity
  - $\hat{y}(d=1) = \arg\min_a ||r - \Phi a||_\infty$
- Constraint set on $\hat{y}(d=1)$
  - Addition profiles monotonically non-decreasing
  - Attrition profile monotonically non-increasing
  - Maximum value that is also optimized
  - Slope constraint
- Inverse problem or deconvolution problem

Productivity profile estimation headcount data

Productivity profile estimation results


(1.67 times larger ℓ∞ error)
New seller productivity discussion

- Some different insights provided by data analysis than thought by leaders in the sales organization
  - No ramp up time for new sellers with experience
  - Longer period of zero productivity for new sellers from universities
  - Slower ramp up for transfers from other business units
  - Attrited sellers have a residual effect
The other examples

- Example 5: product recommendation
  - Collaborative filtering
  - Matrix factorization

- Example 6: talent retention (percentage away from average salary)
  - Kernel regression

- Example 6b: talent retention (mere act of being promoted or being given salary raise)
  - Similar to new hire productivity deconvolution problem

- Example 7: post-disaster philanthropic giving
  - Combination of techniques similar to business impact of outsourcing estimation and new hire productivity estimation

- All of the problems can benefit from the decision tree partitioning of attribute space to obtain groups of objects with similar dose-response behavior
Predictive analytics vs. prescriptive analytics

- All of the dose-response problems presented thus far have been about predictive analytics
- Estimate relationships seen in historical data that will generalize to the future
- This set of predictive analytics provides nice and clear insights to decision makers, but does not go one step further and give actual concrete actions for the decision maker to take
- Giving actions moves from the realm of predictive analytics to the realm of prescriptive analytics
- Predictive analytics: estimation
- Prescriptive analytics: estimation + optimization
- Estimation is from historical data, but optimization is applied to the current data (making use of relationships learned from estimation)
Sales team job role mix prescriptive analytics

- Client attributes $x_i$
- Dose $d_i \in [0, \infty)$
- Transactional revenue from client $r_i \in [0, \infty)$
- Sales expense at client $e_i \in [0, \infty)$
- Aspirational revenue from client $a_i \in [0, \infty)$
- Regard $(r_i - e_i)$ as profit
- Response $y_i = (r_i - e_i)/a_i$

Kernel regression along with decision tree partitioning of attribute space gives us:
- Dose-response functions for different groups of clients: $y^{(k)}(d), k = 1, \ldots, m$
- Standard deviation around dose-response functions: $\sigma^{(k)}(d), k = 1, \ldots, m$

Pose a constrained optimization problem whose objective is to maximize total expected profit achieved from all clients
Constraints

- Constrain the total variance of profit over all clients
  - minimize risk
  - similar to portfolio approach in investing (Markowitz)
  - also helps reduce uncertainty around individual client performance
- Business staffing constraints
  - Exactly same sellers available or very little perturbation of overall salesforce allowed
  - Sellers cannot cross region boundaries
  - Other similar constraints
Job role mix optimization problem

- Given a portfolio of new clients with attributes \( \{x_1, \ldots, x_n\} \) to indicate membership in one of the \( m \) groups and aspirational revenue amounts \( \{a_1, \ldots, a_n\} \), find the optimal \( \{d_1, \ldots, d_n\} \)
  - New clients indicates that this is not necessarily exactly the same set of clients from whom \( y^{(k)}(d) \) and \( \sigma^{(k)}(d) \) were learned
  - The aspirational revenue amounts for the new time period will certainly not be the same as in the historical data used in estimation

- Maximize: \( \Sigma_{i=1,\ldots,n} a_i y^{(k[i])}(d_i) \)

- Subject to: \( \Sigma_{i=1,\ldots,n} a_i^2 \sigma^{(k[i])}(d_i)^2 \leq C \)

- Also subject to certain additional business constraints on \( \{d_1, \ldots, d_n\} \) being satisfied
Talent retention prescriptive analytics

- Assume that a business has allocated an extra $C$ dollars to spend on salary raises this year.
- It wants to spend that money wisely to keep costs of goods sold low.
- Each employee has an original probability of voluntarily attrition $y^o_i$ at some point during the year and an original annual salary $s^o_i$.
- If the employee leaves, then the business will have to hire a replacement employee.
  - Hiring incurs a fixed onboarding cost $c_i$, which includes costs of interviewing, training, overhead, etc.
  - In most instances, a newly hired employee with the same capabilities as the employee he or she is replacing commands a higher salary due to market conditions:
    - Known as salary premium.
    - Incurred for the time period during the year that the replacement is employed.
    - Denote annual salary of replacement as $s'^i$.
- A salary raise of $r^i_i \geq 0$ is offered to each employee at the beginning of the year.
- The constrained optimization problem is to determine the optimal $r^i_i$ subject to $\sum_{i=1}^{n} r^i_i \leq C$. 
Talent retention prescriptive analytics (continued)

- There is a monotonic mapping between the employee’s salary amount and the dose (percentage away from average salary of peers)
  - We can thus write the dose-response function as a function of salary amount if we specify it to the employee: \( \hat{y}_i(\text{indicator of raise; time; salary}) \)
  - Original probability \( \hat{y}_i(0; t; s_{io}) \)
  - Probability after non-zero raise \( \hat{y}_i(1; t; s_{io} + r_i) \)

- The problem is to find the \( r_i \) that minimize the following objective:

\[
\sum_{i=1}^{n} (s_{io} + r_i) \int_{t=0}^{1} 1 - \hat{y}_i(\mathbb{1}[r_i > 0]; t; s_{io} + r_i) \, dt + (s_{io}^r + c_i) \int_{t=0}^{1} \hat{y}_i(\mathbb{1}[r_i > 0]; t; s_{io} + r_i) \, dt
\]

- Subject to: \( r_i \geq 0, i = 1, \ldots, n \) and \( \Sigma_{i=1}^{n} r_i \leq C \)

- In the business setting, it would make sense to only offer discrete raises such as either no raise, a 5% raise, a 7% raise, or a 10% raise

- Such an integer optimization problem can be mapped to the multidimensional knapsack problem and be solved approximately using a greedy algorithm
Dose-response section summary

- Many different business applications can be formulated as dose-response problems
- It is a nice way to frame the problem, because at the end of the day, the recipients of analytics solutions are decision makers who need to decide on doses to achieve favorable outcomes or responses
- Only some variables are *levers* that the decision maker can pull
  - These are dose variables
- Other variables are attributes that cannot be changed by the decision maker, but do mediate the effect of the dose on the response, and thus need to be taken into account in estimation
  - They also provide actionable subdivisions of overall populations
- Response variables should tie back to business goals such as return on assets and return on equity
- Estimation of dose-response relationships is the first step in an overall analytics solution
- To provide even more value by moving to prescriptive analytics, the estimation can and should be coupled with optimization
Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems

**Binary classification problems**

Relational and graph-based problems
Quantile problems
What next?
Binary classification problems

- Of course, not every business analytics problem is a dose-response problem
- Another class of problems are binary classification problems
- From feature vectors $\mathbf{x} \in \mathcal{X}$ predict class label $y \in \{+1, -1\}$
- Learn classifier decision rule $\hat{y}(\mathbf{x})$ from labeled training data $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ and apply to new unseen samples $\mathbf{x}$ from the same distribution
- Variety of algorithms with different strengths and weaknesses
  - Margin-based classifiers including logistic regression and support vector machines
  - Decision trees including CART, CHAID, C5.0
  - Decision lists
  - Nearest neighbor
- Business analytics requirements
  - Ability to deal with categorical features
  - Ability to deal with missing data
  - Interpretability of decision rules
  - Confidence value output along with predicted label
Binary classification examples

- Predict which existing outsourcing clients will reduce or rescope their contracts
  - Key financial indicators
  - Client satisfaction surveys
  - Contract details
  - Significant developments such as change in chief executive officer or restructuring
  - Previous history of contract changes

- Predict salespeople who will not be able to meet their sales quota
  - Sales pipeline data
  - History of quota attainment in previous periods
  - To-date quota attainment in current period
  - Human resources data

- Simply predict the individuals who will voluntarily attrit
  - Compensation
  - Time since hiring, since promotion, since salary increase
  - Performance ratings
  - Business unit and region
Example of attrition modeling C5.0 decision tree decision rules

- Rules that predict a salesperson will voluntary resign
  - Rule 1
    - Job Role = Technical Seller
    - Months Since Promoted > 13
    - Months Since Hired ≤ 99
    - Business Unit = New England
  - Rule 2
    - Job Role = Specialty Seller
    - Base Salary ≤ $75,168
    - Months Since Promoted > 13
    - Months Since Promoted ≤ 30
    - Quota-Based Compensation Plan = False
Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems
What next?
Relations

- Much of business revolves around who you know and what you know
- 2012 is being called ‘the year of social business’
- Connections and networking
- All sorts of links
  - Companies connected to other companies
  - People connected to other people
  - People connected to companies
  - Companies connected to products
  - People connected to products
  - People connected to knowledge
  - Knowledge connected to products
A social network of companies

- Various types of edges connect companies, ranging from board members having common memberships, to similar strategies such as early technology adopter or environmentally conscious, to business partnerships.

- The social network can be visually displayed to help marketing and sales understand the risks and opportunities among a certain set of accounts.

Highlight companies that are not currently customers (whitespace) that could become customers

“The Dark Side”
Competitive zone of influence
Identification of sales experts within a company

- Enterprise + market social network: which sellers should staff specific accounts to yield higher revenues
- Can also have products as a third type of node

Diagram:
- Seller
- Client
- Relationship

seller-client relationships
client-client relationships
seller-seller relationships
Company acquisition

- Which acquisition target will expand the company’s customer base into whitespace

customer node size indicates revenue
Company acquisition (continued)

- Which acquisition is better?

<table>
<thead>
<tr>
<th>Company</th>
<th>Customer</th>
<th>Whitespace</th>
<th>Potential Acquisition</th>
<th>Competitor</th>
<th>Competitive Account</th>
<th>Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A has good whitespace coverage. It also sells to many of our clients. Its whitespace accounts have similar characteristics to existing IBM accounts. This indicates that A's whitespace accounts could have high propensity to buy other products/services from IBM.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B has good whitespace coverage. It also sells to many of competitor's clients. B could be the acquisition of choice if we are considering penetrating into competitive accounts.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C has good whitespace coverage. Its clients however, does not seem to be buying any other products (IBM or competitive) indicating low propensity to grow revenue in other businesses.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D has almost no whitespace coverage, primarily serving several of our key accounts.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

customer node size indicates revenue
What pieces of marketing collateral should be recommended to sellers?

Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems
What next?
Quantile problems

- Quantiles are equal divisions of the cumulative distribution function of a random variable
  - 2-quantile is median
  - 100-quantiles are the percentiles
  - $k^{th}$ $q$-quantile of random variable $X$ with sample value $x$ is $\Pr[X < x] \leq k/q$

- Useful in dividing data points into subsets of equal cardinality

- Quantile regression is the estimation of quantiles from data
  - Does not have as nice a geometry as least-squares
  - Requires linear programming, etc.

- Quantile-based analysis arises in business analytics in different contexts
Revenue opportunity estimation

- Revenue opportunity or wallet
  - Total funds a customer has allocated to spend on a certain category of products
  - Greatly affects what and how much they purchase
  - Not known to vendors, but would be useful for sales strategy and customer targeting

- Wallet estimation using quantile regression
  - Wallet amount is a hidden variable
  - Demographics/firmographics are observed variables
  - Historical spending is an observed variable

- Customer spends less than or equal to wallet amount

- Customers rarely spend all of their money at one vendor
  - Realistic wallet amount is a high quantile, e.g. 80th percentile of spending distribution

- Estimation procedure is conditional quantile regression of spending as a function of demographic/firmographic features from collection of historical samples

Giving scores or grades to salespeople

- Example of descriptive analytics
- Grade every seller so that good sellers have high scores and bad sellers have low scores
- Then can aggregate scores to see which business units or regions are divisions of strength or weakness within the overall salesforce
- Theory of grading in education is predicated on population following bell curve
  - Not everything is Gaussian
- Exponential and Pareto distributions
  - Sports analytics
    - Since professional athletes are the best of the best, their distribution is at the tail of the general population (which would be Gaussian)
  - Economics
    - Allocation of wealth among individuals
  - Revenue achieved by and quota attainment fraction of salespeople
Pareto and exponential distributions well-describe quota attainment
Quantile-based scoring

- With such a distribution, scoring like this obviously doesn’t make sense:

- Scoring based on percentiles
- Simple formula to calculate percentile-based cutoffs with exponential distribution model
  - Cutoff = $-\mu \ln(1 - p)$, where $\mu$ is the sample average and $p$ is the percentile

Source: Wikimedia Commons
Agenda

What is business analytics
Signals and systems view of a business
A detailed example: sales team job role mix
Perspectives on business analytics
Dose-response problems
Binary classification problems
Relational and graph-based problems
Quantile problems

What next?
Summary

- Business analytics is an important endeavor that is behind the transformation of many successful companies.
- A business can be thought of as a feedback system with funds as the signal:
  - Desire some amount of positive feedback
  - Return on assets and return on equity greater than one
- A wide variety of business problems can be approached through statistical signal processing techniques:
  - Estimation theory is really the core subject of predictive analytics
- An important way to view business analytics problems is through dose-response formulations
- Other types of formulations are also important in descriptive, predictive, and prescriptive analytics
- Graph-based methodologies are becoming more and more important in this ‘year of social business’
- Business analytics is a domain looking for innovative statistical signal processing
Pursuing research in business analytics

- Interesting signal processing problems arise because of the different types of data and signals, their varying relationships to the business questions of interest, and the interestingness of the questions themselves.

- Unfortunately, data is the new currency.

- However, theoretical models distinguish statistical signal processing from data mining:
  - Data mining is data-centric but signal processing is not.

- “Mathematical problems are formulated and motivated by specific classes of applications and by the methods of solution that one already knows. Solutions are produced. Throughout this process, questions are asked: How can we extend existing mathematical methodologies? How can we use existing methodologies in the context of a specific physical problem to obtain a tractable formulation which addresses the issues of interest in the more ill-defined physical problem? Why are we working on these problems? Will they increase the scope of theory or theoretical understanding? Or do they meet the specific needs of a class of physical problems? What types of problems should we be attempting to formulate, motivated either by theory or practice?”

Get involved

- Thus, business analytics research can be pursued even without access to real-world data
  - Of course, it is better to have some for motivation and validation

- Collaboration
  - Companies are starting to awaken to the need for solid, innovative analytics
    - They don’t necessarily know how to get started or what questions can be addressed
    - Opportunity for consulting
  - Businesses are everywhere so lots of possible subject matter expert collaborators
    - From your neighborhood convenience store to large corporations
    - Perhaps you are a business owner yourself
    - Unlike more specific industries like defense, electronics, astronomy, geophysics, etc.
    - Schools of business and management in your university

- Directly influence the ‘bottom line’ of a company
  - Become a business analytics professional for a company

- Exciting new field that is redefining the way the world works
  - Get involved
Acknowledgements

- IBM Research

- IBM Business Units
Thank you and further questions
Works cited


