Creating and facilitating positive revolutionary growth through:

Education  Implementation  Connection

We’re bridging the gap between business and data science.
Here’s what we’re going to do today:

1. Provide clarity
2. Seed your plan
3. Prepare for launch
Terminology matters

Big Data → Machine Learning → Artificial Intelligence

Fuel → Engine → Mechanical Power
Enterprise AI

Automation

next-generation opportunities

Analytics

skip-generation opportunities
The primary purpose of automation is to perform a repetitive task previously performed by humans.
Automation is no longer limited to mechanical applications

<table>
<thead>
<tr>
<th>Vision</th>
<th>Speech</th>
<th>Language</th>
<th>Robotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>Visual inspection</td>
<td>Voice call trees</td>
<td>OCR</td>
</tr>
<tr>
<td>Cheaper</td>
<td>Video security</td>
<td>Chatbots</td>
<td>Robotic arms</td>
</tr>
<tr>
<td>Faster</td>
<td>Object ID in image posts</td>
<td>Video transcription</td>
<td>Sentiment analysis</td>
</tr>
<tr>
<td>Better</td>
<td>Radiology</td>
<td>Voice ID</td>
<td>Author ID</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hazardous conditions</td>
</tr>
</tbody>
</table>
Framing the Opportunity: Imagery

If you could identify something specific in photographs, illustrations and/or video, what could you do with that data?

EXAMPLE

If you could identify a candidate’s facial expressions, you could better evaluate his/her interest in the job.
Your turn...

If you could identify ________________ in ________________, you could ________________.

Hint: sources could include anything from social media to surveillance footage to product/service-related image capture, to name a few.
**Framing the Opportunity: Audio**

If you could recognize something specific in live and recorded audio sources, what could you do with that data?

**EXAMPLE**

If you could recognize a customer’s emotional state in his/her tone of voice, you could customize our response.
Your turn...

If you could recognize _________________ in _________________, you could _________________.

Hint: audio sources could include data you capture (customer service interactions) as well as data you consume (podcasts).
Framing the Opportunity: Language

If you could interpret something in the language you generate and consume, what could you do with that data?

EXAMPLE

If you could interpret sentiment in the claims forms you process, you could minimize settlement pay-outs.
Your turn...

If you could interpret ________________ in ________________, you could _____________________.

Hint: consider focusing on one source, something your company, customers, competitors or outside content producers produce.
Enterprise AI

Automation
next-generation opportunities

Analytics
skip-generation opportunities
Sample Data

Algorithm

New Data

Model

Prediction
Training Data

Algorithm

New Data

Machine Learning Model

Prediction
The Humble Regression Model

Femur Length vs Height (in)

Six Segment Comparison

- Mule African
- Male Asian
- Male European
- Female African
- Female Asian
- Female European

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Sales Regression Model

Q4 Forecast

US $
What’s changed?

• Hyper-connected world makes 100% sampling feasible
• Massive amounts of data on each individual customer
• We can predict their behavior without even asking
Break
At the heart of every AI application is a machine learning model.
New Data

Algorithm

Machine Learning Model

Prediction

1 Number
2 Probability
3 Category
Number Models

• Similar to “traditional” regression models, just much more data

• Useful for predicting an exact quantity in advance

• Can be any continuous unit of measure: $, time, units, degrees
New Data

Number Model

Forecasting

Historical
Quarterly Sales

<table>
<thead>
<tr>
<th>Demographics, interests, purchase history, ...</th>
<th>$ per quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 1</td>
<td>$557</td>
</tr>
<tr>
<td>Customer 2</td>
<td>$667</td>
</tr>
<tr>
<td>Customer 3</td>
<td>$223</td>
</tr>
<tr>
<td>Customer n</td>
<td>$1009</td>
</tr>
</tbody>
</table>

Algorithm

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon \]

Prediction

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Historical Customer Lifetime Value

<table>
<thead>
<tr>
<th>Demographics, interests, purchase history, ...</th>
<th>Lifetime spend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 1</td>
<td>$3556</td>
</tr>
<tr>
<td>Customer 2</td>
<td>$456</td>
</tr>
<tr>
<td>Customer 3</td>
<td>$1225</td>
</tr>
<tr>
<td>Customer n</td>
<td>$3778</td>
</tr>
</tbody>
</table>

Algorithm

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_n X_n + \epsilon \]

Prediction

New customer
### Historical Maximum Price

<table>
<thead>
<tr>
<th>Demographics, interests, purchase history, ...</th>
<th>Max Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 1</td>
<td>$23.00</td>
</tr>
<tr>
<td>Customer 2</td>
<td>$14.00</td>
</tr>
<tr>
<td>Customer 3</td>
<td>$38.00</td>
</tr>
<tr>
<td>Customer n</td>
<td>$26.00</td>
</tr>
</tbody>
</table>

**Algorithm**

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon \]

**Prediction**
Historical Segment Times

<table>
<thead>
<tr>
<th>Segment</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>n</td>
<td>8</td>
</tr>
</tbody>
</table>

Algorithm

New request

Machine Learning Model

Prediction

Number Model Process Optimization

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Framing the Opportunity: Forecasting

If we could predict precisely how much each guest will spend in our casino over the lifetime of our relationship, we could determine how much effort to put into maintaining the relationship with each guest.

If we could predict precisely the maximum amount each customer would spend, we could maximize ticket and package prices.
Your turn...

If you could predict precisely how much/many ______________________ at any given moment, you could _________________________________.

Hint: identify places where you currently rely on estimates and averages to make decisions, then imagine instead that you have information about each individual.
Framing the Opportunity: Optimization

If we could identify the fastest way to get from one place to another at a given time, we could save both time and relationships.

If we could identify the fastest way to route patients through the clinic at any given time, we could maximize the number of patients seen.
Your turn...

If you could identify the fastest way to ___________________________,
you could __________________________________________________.

Hint: think about processes in your business that require moving something from point A to point B with multiple paths to choose from, then imagine being able to predict the fastest path at any given time.
Probability Models

• Commonly referred to as recommenders or recommendation engines

• Predict the probability of a given event occurring (e.g. purchase)

• Useful when you have two long lists and need to make connections
People

Mary
John
Fred
Byron
Nick
Spencer
David
Doug
Stacy
Constance
Robert
Darcy

Movies

Mary Poppins
Jaws
The Shining
The Sound of Music
Get Out
Blazing Saddles
Chinatown
Die Hard
The Godfather
50 Shades of Grey
Slumdog Millionaire
Halloween

Demographics
Purchase history
Hobbies and interests
etc.
Probability Model
Personalization

<table>
<thead>
<tr>
<th>Mary Poppins</th>
<th>Jaws</th>
<th>The Shining</th>
<th>The Sound of Music</th>
<th>Get Out</th>
<th>Blazing Saddles</th>
<th>Chinatown</th>
<th>Die Hard</th>
<th>The Godfather</th>
<th>50 Shades of Grey</th>
<th>Slumdog Millionaire</th>
<th>Halloween</th>
</tr>
</thead>
<tbody>
<tr>
<td>% % % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % % % %</td>
</tr>
<tr>
<td>John</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Fred</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Byron</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Nick</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Spencer</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>David</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Doug</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Stacy</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Constance</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Robert</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
<tr>
<td>Darcy</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % % % % %</td>
<td>% % %</td>
<td>% % % % %</td>
<td>% % % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
<td>% % %</td>
</tr>
</tbody>
</table>
### Historical Recommendation Outcomes

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Customer info + movie info</th>
<th>Purchased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation 1</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Recommendation 2</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Recommendation 3</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Recommendation n</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

---

#### Algorithm

New login → Machine Learning Model → Prediction

#### Probability Model

Personalization

---

Taming Dragons: A breakthrough approach to AI for business leaders

copyright: milk+honey

milkandhoney.ai

hello@milkandhoney.ai

Presented at the O'Reilly AI Conference in San Francisco

September 5, 2018
Probability Model
Advanced Expert Systems

Response
Recommenders

Issues

Responses

Claims

Arguments

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Guided Selling

SIEMENS

Next action

Sale status

Probability Model

Advanced Expert Systems

Training Coach

SIOP

Goal & status

Next exercise

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Framing the Opportunity: Personalization

If we knew which **movies** were most likely to be viewed by each individual **user**, we could **maximize streaming revenue**.

If we knew which **articles** were most likely to be read by each individual **reader**, we could **increase reader satisfaction**.
Your turn...

If you knew which ____________________________ were most likely to ____________________________ by each individual ____________________________, you could ____________________________.

Hint: identify places where you use segments or groupings, and imagine being able to engage each individual directly.
Framing the Opportunity: Prioritization

If we could rank sales leads by their likelihood to convert we could focus on the most lucrative deals.
Your turn...

If you could rank _____________ by their likelihood to ______________, you could ____________________________________.

Hint: think of areas in your business where it’s difficult to decide what’s most important, or where to focus resources, and imagine eliminating that guesswork.
Framing the Opportunity: Advanced Expert Systems

If we could match incoming calls with proven resolution responses, we could improve customer service.

If we could match client requirements with ideal solution recommendations, we could reduce RFP response times.
Framing the Opportunity: Advanced Expert Systems

If we could match incoming calls with proven resolution responses, we could improve customer service.

If we could match client requirements with ideal solution recommendations, we could reduce RFP response times.
Category Models

• Commonly referred to as classifiers

• Used to classify elements of a population into one of two or more categories

• Binomial (two category) models used for “yes/no” type decisions

• Multinominal (three more categories) used for selection among options
Hi,

Website Administrator

SEO is the best way to increase your business volume. And we have an expert SEO team.

We can quickly promote your website. We can place your website on top of the Natural Listings on Google, Yahoo and MSN.

→ Position your website to be top ranking
→ Refine your website design to be engaging
→ Increase profitability click-through rates from PPC campaigns
→ Develop strong conversion rates
→ Expert web statistics analysis

Our prices are less than half of what other companies charge.

We would be happy to send you a proposal using the top search phrases for your area of expertise.

Note: Please must check our past record and current client status.

WE ALSO DESIGN & DEVELOPE THE WEB-SITE FLASH PHP, Joomla, Open source, E-Commerce at reasonable cost.

Thanks & Regards

Mark
mark.seotech@gmail.com

SEO - Link Building - Copyrighting - Web Designing - PHP

Date (1 day ago) ★

Hi, Mark Mark

What? You don’t have your own domain?

Bolding it doesn’t make it my name

You have what?

Do what?

Must I? OK Then.

Are you shouting now?

Less than 1/2 the cost of your competitors apparently

I think you mean copyright.
### Email Examples

<table>
<thead>
<tr>
<th>Misspellings, grammatical errors, all caps, domain...</th>
<th>Spam?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email 1</td>
<td>yes</td>
</tr>
<tr>
<td>Email 2</td>
<td>no</td>
</tr>
<tr>
<td>Email 3</td>
<td>yes</td>
</tr>
<tr>
<td>Email n</td>
<td>no</td>
</tr>
</tbody>
</table>

### Algorithm

- **New email**
- **Machine Learning Model**
  - If > 80% then spam
  - If <80% then not spam
- **Prediction**
  - Model calculates the probability that it is spam
  - Uses rule YOU set to decide the final output

### Category Model

- **Binomial Detection**
  - 100% Spam
  - 80% Not spam
  - 0% Not spam

---

**Taming Dragons: A breakthrough approach to AI for business leaders**

copyright: milk+honey

Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
Previous Patient Diagnosis

<table>
<thead>
<tr>
<th>Symptoms, medical history,...</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>Eczema</td>
</tr>
<tr>
<td>Patient 2</td>
<td>Allergy</td>
</tr>
<tr>
<td>Patient 3</td>
<td>Eczema</td>
</tr>
<tr>
<td>Patient n</td>
<td>Psoriasis</td>
</tr>
</tbody>
</table>

Algorithm

- Model calculates the probability for each possibility
- Selects the most likely

Machine Learning Model

<table>
<thead>
<tr>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eczema</td>
</tr>
<tr>
<td>Allergy</td>
</tr>
<tr>
<td>Psoriasis</td>
</tr>
</tbody>
</table>

Category Model
Multinomial Detection
Historical Loan Outcomes

<table>
<thead>
<tr>
<th>Loan</th>
<th>Customer + loan details, ...</th>
<th>Paid?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan 1</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Loan 2</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Loan 3</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Loan n</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Algorithm

Machine Learning Model

- 85%
- If > 90% then approve
- If <10% then decline

Category Model
Binomial Prediction

85% If > 90% then approve
If <10% then decline
Framing the Opportunity: Binomial Detection

If we could detect deceptive, distracting, unproductive, annoying, aggravating communication, from historical communication data we could the good stuff.
Your turn...

If you could detect __________________________ from __________________________ data, you could __________________________.

Hint: look for places where you make pass/fail type decisions in your current business processes.
Framing the Opportunity: Binomial Prediction

If we could predict the likelihood that a loan applicant will default on his/her loan from historical loan data, we could let ‘em have it anyway.
Your turn...

If you could predict ___________________________ from ___________________________ data, you could ___________________________.

Hint: look for places where you make yes/no type decisions in your current business processes.
Map the opportunities you identified to your company’s functional areas…

Your Company

- Logistics & Supply Chain
- Marketing & Promotion
- Product/Service Development
- Production/Manufacturing
- Research & Development
- Sales
- Accounting & Finance
- Administration & Operations
- Customer Service Support
- Human Resources
- Legal
You are here, sort of

Taming Dragons 9/5/18
Opportunity Assessment
Strategic Plan
PMO & go
Coordinated Transformation

Discover → Design → Develop → Deploy

- Machine Learning
- Data & Platform
- Corporate Culture
- Workforce
Machine Learning

• Determine potential ROI for each application
• Conduct feasibility assessments
• Estimate model development times
• Create staged roadmap
Data & Platform

- Select data science platform
- Create end-state architecture
- Determine implementation plan synchronized with ML roadmap
Corporate Culture

• Unprecedented teamwork
• Data-driven decision making
• Culture of experimentation
• C-suite commitment
Workforce

• Upskilling plan for existing staff
• Reskilling forecast & alternatives for displaced workers
• Forecast, acquire and grow data science talent from within
Scarcity of data science talent?

Open jobs in US for “Data Scientist”

19k

Data Scientists in US “actively” seeking

15k

Data Scientists in US “quietly” seeking

22k

Sources: LinkedIn 8/7/18
Presented at the O'Reilly AI Conference in San Francisco
September 5, 2018
## Candidate Profile

**Mary Smith**

### Development Role History

<table>
<thead>
<tr>
<th></th>
<th>Business Objective</th>
<th>Solution Design</th>
<th>Model Creation</th>
<th>Model Evaluation</th>
<th>Model Deployment</th>
<th>Continuous Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Machine Learning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data &amp; Platform</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Skills

#### Technical

- **Scripting languages**
  - Python
  - R
- **Math Computation Tools**
  - MATLAB

© milkandhoney 2018
## Data Science Project Portfolio

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Project duration</th>
<th>Project type</th>
<th>ML application</th>
<th>Functional Business Area</th>
<th>ML implementation</th>
<th>Algorithm family</th>
<th>Level of contribution</th>
<th>Project scoping</th>
<th>ML application design</th>
<th>Model development</th>
<th>Model evaluation</th>
<th>Model deployment</th>
<th>Model evolution</th>
<th>Data platform architecture</th>
<th>Data integration</th>
<th>Data platform development</th>
<th>Data visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>207 BCW Topic</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier -</td>
<td>regression</td>
<td>Classification -</td>
<td>regression</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>207 Mushroom PCA</td>
<td>1-3 months</td>
<td>class</td>
<td>Classification</td>
<td>regression</td>
<td>Classification -</td>
<td>regression</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>207 Final Project</td>
<td>1-3 months</td>
<td>class</td>
<td>Classification</td>
<td>bayesian</td>
<td>Classification -</td>
<td>bayesian</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>201 email classification</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier -</td>
<td>bayesian</td>
<td>Classification -</td>
<td>bayesian</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>201 spam detector</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier -</td>
<td>bayesian</td>
<td>Classification -</td>
<td>bayesian</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>201 product</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommendation</td>
<td>regression</td>
<td>Recommendation -</td>
<td>regression</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>201 product</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommendation</td>
<td>regression</td>
<td>Recommendation -</td>
<td>bayesian</td>
<td>Team member</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Taming Dragons 9/5/18

Opportunity Assessment

Strategic Plan

PMO & go
PMO

• Need to share learnings, encourage teamwork
• Can’t centralize because too domain dependent
• Moving toward multidisciplinary teams
We’re a go!

- Taming Dragons
- Opportunity Assessment
- Strategic Plan
- PMO & go

Taming Dragons 9/5/18
Once you get a win like this, generating real revenue and increasing profitability, it really opens everyone’s eyes to where else could we use machine learning technologies to enhance the business.

Ken O’Brien, CIO - RR Donnelley
Q&A
Automation Exercises

Framing the Opportunity: Imagery

If I could identify ________________________________ in ________________________________,
I could ________________________________.

Hint: sources could include anything from social media to surveillance footage to product/service-related image capture, to name a few.

Framing the Opportunity: Audio

If I could recognize ________________________________ in ________________________________,
I could ________________________________.

Hint: audio sources could include data you capture (customer service interactions) as well as data you consume (podcasts).

Framing the Opportunity: Language

If I could interpret ________________________________ in ________________________________,
I could ________________________________.

Hint: consider focusing on one source, something your company, customers, competitors or outside content producers produce.
Number Model Exercises

Forecasting
If I could predict precisely how much/many ___________________________ at any given moment, I could ___________________________.

Hint: identify places where you currently rely on estimates and averages to make decisions, then imagine that instead you have information about each individual.

Optimization
If I could identify the fastest way to ___________________________,
I could ___________________________.

Hint: think about processes in your business that require moving something from point A to point B with multiple paths to choose from, then imagine being able to predict the fastest path at any given time.
Probability Model Exercises

**Personalization**
If I knew which were most likely to be used by each individual, I could engage each individual directly.

Hint: identify places where you use segments or groupings, and imagine being able to engage each individual directly.

**Prioritization**
If I could rank by their likelihood, I could prioritize.

Hint: think of areas in your business where it’s difficult to decide what’s most important, or where to focus resources, and imagine eliminating that guesswork.

**Advanced Expert Systems**
If I could match with, I could leverage.

Hint: think of opportunities to more effectively and immediately leverage your knowledge base to expedite and improve process.

**Category Model Exercises**

**Binomial Detection**
If I could detect from, I could improve.

...
I could _________________________________.

Hint: look for places where you make pass/fail type decisions in your current business processes.

**Binomial Prediction**

If I could predict _________________________________
from _________________________________,
I could _________________________________.

Hint: look for places where you make yes/no type decisions in your business today.
Additional Application Exercise Examples

**Automation**
I would love to be able to recognize faces in images from surveillance cameras in order to improve security at our factories.

I would love to be able to interpret customer emotional state in audio from chatbot conversations in order to improve customer relations.

I would love to be able to identify prospective customers in text from blog posts in order to increase sales.

**Number Models**
If I could predict precisely how long terminally ill patients have to live at any given time, I could better provide them with the care they need.

If I could predict precisely how much electricity was going to be demanded at any given moment I could reduce costs by optimizing the power grid.

If I could predict precisely how many more months each engineer was going to stay at any given time I could start sourcing replacements sooner.

**Probability Models**
I would love to know which candidates are the most likely to be successful for each individual job opening in order to minimize unnecessary interviews.

I would love to be able to rank outstanding repair orders by level of potential damage in order to minimize risk of injury.

I would love to be able to capture and share product configuration expertise with operators in our call center in order to maximize productivity.

**Category Models**
I would love to be able to predict whether middle school students are at low, medium or high risk of dropping out from school record data in order to focus counseling resources appropriately.

I would love to be able to predict which suicide hotline callers will actually attempt suicide from their call history and current emotional disposition data in order to initiate actions to prevent it.

I would love to be able to detect children that have been victims of abuse from school and health record data in order to provide them with early mental health care.
Data Science Candidate Profile
ZR_520_CAND

Candidate Overview
Current Title: Director of Data Science
Work Experience (yrs): 15
Degrees held: Master of Science; Bachelor of Science

Development Role History

Tool & Technology Skills

Machine Learning Application Experience

9/4/2018
## Training & Education Summary

<table>
<thead>
<tr>
<th>Subject</th>
<th>Level of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear algebra</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Calculus</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Basic statistics</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Advanced statistics</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Data visualization</td>
<td>Individual online course</td>
</tr>
<tr>
<td>Basic machine learning</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Advanced machine learning</td>
<td>Accredited degree program</td>
</tr>
<tr>
<td>Deep learning</td>
<td>Bootcamp or certificate program</td>
</tr>
<tr>
<td>Reinforcement Learning</td>
<td>Employer provided training</td>
</tr>
<tr>
<td>Data privacy/governance</td>
<td>n/a</td>
</tr>
<tr>
<td>Natural language processing</td>
<td>Bootcamp or certificate program</td>
</tr>
<tr>
<td>Image processing/computer vision</td>
<td>n/a</td>
</tr>
<tr>
<td>Voice &amp; audio processing</td>
<td>n/a</td>
</tr>
<tr>
<td>Other</td>
<td>n/a</td>
</tr>
</tbody>
</table>

## Machine Learning Experience

<table>
<thead>
<tr>
<th>Algorithm family</th>
<th>Numeric Prediction</th>
<th>Recommendation</th>
<th>Classification</th>
<th>Level of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>item</td>
<td>action</td>
<td>sequential</td>
</tr>
<tr>
<td>Association Rule</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborative Filtering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content Filtering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hybrid Filtering</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decision Tree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep Learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimensionality Reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ensemble</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instance-based (kNN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matrix Factorization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neural Network</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reinforcement Learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Awards & Accomplishments

---

9/4/2018
# Data Science Candidates Profile

## ZR_520_CAND

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Project Duration</th>
<th>Project Type</th>
<th>ML Application</th>
<th>Functional Business Area</th>
<th>ML Implementation</th>
<th>Algorithm Family</th>
<th>Level of Contribution</th>
<th>Project Scoping</th>
<th>ML Application design</th>
<th>Model Development</th>
<th>Model Evaluation</th>
<th>Model Deployment</th>
<th>Model Evolution</th>
<th>Data Platform Architecture</th>
<th>Data Integration Plan</th>
<th>Data Platform Development</th>
<th>Data Visualization</th>
<th>Model Automation</th>
<th>Data Governance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOW Topic Classification</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Marketing &amp; promotion</td>
<td>Classification - natural language</td>
<td>Regression</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mushroom PCA &amp; clustering</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Marketing &amp; promotion</td>
<td>Classification - natural language</td>
<td>Regression</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Project</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - image</td>
<td>Product/service development</td>
<td>Classification - image</td>
<td>Deep Learning</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>email classification</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Administration &amp; operations</td>
<td>Classification - natural language</td>
<td>Bayesian</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>spam detector</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Administration &amp; operations</td>
<td>Classification - natural language</td>
<td>Bayesian</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>product recommender</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommend - er</td>
<td>Marketing &amp; promotion</td>
<td>Recommendation - item</td>
<td>Regression</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tweet clustering by source</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Marketing &amp; promotion</td>
<td>Classification - natural language</td>
<td>Instance-based ANN</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>joins, book EDA, synonyms</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - natural</td>
<td>Administration &amp; operations</td>
<td>Classification - natural language</td>
<td>Bayesian</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shortest path wikipedia</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommend - er</td>
<td>Administration &amp; operations</td>
<td>Recommendation - item</td>
<td>Decision Tree</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pagerank</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommend - er</td>
<td>Marketing &amp; promotion</td>
<td>Recommendation - item</td>
<td>Decision Tree</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wordcount &amp; sort with spark</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Customer service</td>
<td>Classification - natural language</td>
<td>Instance-based ANN</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>binary classification</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Customer service support</td>
<td>Classification - natural language</td>
<td>Instance-based ANN</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTR with logistic regression</td>
<td>1-3 months</td>
<td>class</td>
<td>Recommend - er</td>
<td>Logistic &amp; supply chain</td>
<td>Recommendation - item</td>
<td>Regression</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>decision trees</td>
<td>1-3 months</td>
<td>class</td>
<td>Numeric Prediction - item</td>
<td>Marketing &amp; promotion</td>
<td>Numeric prediction</td>
<td>Decision Tree</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adder in tenasflow</td>
<td>1-3 months</td>
<td>class</td>
<td>Numeric Prediction - item</td>
<td>Sales</td>
<td>Numeric prediction</td>
<td>Deep Learning</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>word similarity language model</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - natural</td>
<td>Marketing &amp; promotion</td>
<td>Classification</td>
<td>Deep Learning</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sentence probability</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - natural</td>
<td>Marketing &amp; promotion</td>
<td>Classification</td>
<td>Deep Learning</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS tagging, pipe cutting word id</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - natural</td>
<td>Administration &amp; operations</td>
<td>Classification - natural language</td>
<td>Deep Learning</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fake news detector</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Administration &amp; operations</td>
<td>Classification - natural language</td>
<td>Deep Learning</td>
<td>Team lead</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>election results tweet analysis</td>
<td>1-3 months</td>
<td>class</td>
<td>Classifier - condition</td>
<td>Marketing &amp; promotion</td>
<td>Classification</td>
<td>Regression</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capstone - crisis navigator</td>
<td>3-6 months</td>
<td>class</td>
<td>Automation - natural</td>
<td>Product/service development</td>
<td>Classification - natural language</td>
<td>Bayesian</td>
<td>Team lead</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Classification</td>
<td>1-3 months</td>
<td>class</td>
<td>Automation - image</td>
<td>Administration &amp; operations</td>
<td>Classification - image</td>
<td>Instance-based ANN</td>
<td>Team</td>
<td>member</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Science Job Opening Summary
Acme Tech - Data Science Manager - ZR_36_JOB

Job Opening Overview
Team Role: Manager
Experience Requirement: 5+ years
Education Requirement: Masters

Job Opening Scope
Machine Learning
Data & Platform

Experience Requirement:
5+ years
Machine Learning

Education Requirement:
Masters

Tool & Technology Skills

Development Platforms
- Amazon Web Services
- IBM Cloud
- Google Cloud Platform
- Microsoft Azure
- On-premise/private

OS's
- Unix/Linux
- Windows
- MacOS

Programming Languages
- Python
- R
- Java
- C/C++

Scripting Languages
- Unix Shell
- awk
- Perl
- Julia

Math Computation Tools
- MATLAB
- Octave
- Mathematica

Visualization Tools
- Tableau
- QlikView
- TIBCO Spotfire
- SAS JMP

Python Tools & Libraries
- Jupyter Notebooks
- TensorFlow
- PyTorch
- Hadoop/Hive/Pig
- Spark/Mlib
- Flume
- scikit-learn

Data Mining Tools
- MS Excel Data Mining
- MS SQL Server Mining
- SAS Enterprise Miner
- Oracle Data Mining
- IBM Watson

Statistical Analysis Tools
- SAS Base
- IBM SPSS Statistics
- Oracle R Enterprise
- Stan
- Minitab
- Statistica

Business Intelligence Tools
- SAS
- SAP Business Objects
- Predictive Analytics
- KNIME
- IBM Cognos
- Anoss

ML Productivity Tools
- RapidMiner
- Cloudera
- Amazon ML
- Azure ML Studio
- Orange
- IBM SPSS Modeler
- DataRobot
- Salford Predictive Modeler

Databases
- Oracle SQL
- Microsoft SQL Server
- PostgresSQL
- Other SQL
- MongoDB
- Amazon DynamoDB
- Couchbase
- Other: NoSQL

Project Details

<table>
<thead>
<tr>
<th>Functional Business Area</th>
<th>Automation</th>
<th>Numeric Prediction</th>
<th>Recommender</th>
<th>Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>image processing</td>
<td>forecasting/ projection</td>
<td>personalization</td>
<td>condition detection</td>
</tr>
<tr>
<td>Accounting &amp; Finance</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Administration &amp; operations</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Customer service support</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Human resources</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Legal</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Logistics &amp; supply chain</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Marketing &amp; promotion</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Product/service development</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Production/manufacturing</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Research &amp; development</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Sales</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Other</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
</tbody>
</table>

9/4/2018
**Candidate Overview:** ZR_520_CAND

**Current Title:** Director of Data Science

**Work Experience (yrs):** 15

**Degrees held:** Master of Science; Bachelor of Science

---

### Development Role Comparison

<table>
<thead>
<tr>
<th>Business Objective</th>
<th>Solution Design</th>
<th>Model Creation</th>
<th>Model Evaluation</th>
<th>Model Deployment</th>
<th>Continuous Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data &amp; Platform</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

### Tool & Technology Skills Match Assessment

- **Development Platforms**
  - Amazon Web Services
  - IBM Cloud
  - Google Cloud Platform
  - Microsoft Azure
  - On-premise/private

- **OS's**
  - Unix/Linux
  - Windows
  - MacOS

- **Programming Languages**
  - Python
  - R
  - Java
  - C/C++

- **Scripting Languages**
  - Unix Shell
  - awk
  - Perl
  - Julia

- **Math Computation Tools**
  - MATLAB
  - Octave
  - Mathematica

- **Visualization Tools**
  - Tableau
  - QlikView
  - TibCO Spotfire
  - SAS JMP

- **Python Tools & Libraries**
  - Jupyter Notebooks
  - TensorFlow
  - PyTorch
  - Hadoop/Hive/Pig
  - Spark/Mlib
  - Flume
  - scikit-learn

- **Data Mining Tools**
  - MS Excel Data Mining
  - MS SQL Server Mining
  - SAS Enterprise Miner
  - Oracle Data Mining
  - IBM Watson

- **Statistical Analysis Tools**
  - IBM SPSS Statistics
  - Oracle R Enterprise
  - Stan
  - Miniconda
  - Statistica

- **Business Intelligence Tools**
  - SAS
  - SAP Business Objects
  - Predictive Analytics
  - KNIME
  - IBM Cognos
  - Angoss

- **ML Productivity Tools**
  - RapidMiner
  - Cloudera
  - Amazon ML
  - Azure ML Studio
  - Orange
  - IBM SPSS Modeler
  - DataRobot
  - Salford Predictive Modeler

- **Databases**
  - Oracle SQL
  - Microsoft SQL Server
  - PostgreSQL
  - Other SQL
  - MongoDB
  - Amazon DynamoDB
  - Couchbase
  - Other NoSQL

---

### Application Experience Comparison

<table>
<thead>
<tr>
<th>Functional Business Area</th>
<th>Automation</th>
<th>Natural Language Processing</th>
<th>Robotics/Industrial Control</th>
<th>Numeric Prediction</th>
<th>Recommendation</th>
<th>Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting &amp; finance</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Administration &amp; operations</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Customer service support</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Human resources</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Legal</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Logistics &amp; supply chain</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Marketing &amp; promotion</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Product/Service development</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Production/Manufacturing</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Research &amp; development</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Sales</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Other/unspecified</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
</tbody>
</table>

---

**Job Opening Candidate Match Summary**

Acme Tech - Data Science Manager - ZR_36_JOB

**Candidate Overview:** ZR_520_CAND

**Development Role Comparison**

<table>
<thead>
<tr>
<th>Experience in role</th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data &amp; Platform</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Tool & Technology Skills Match Assessment**

- **Development Platforms**
  - Amazon Web Services
  - IBM Cloud
  - Google Cloud Platform
  - Microsoft Azure
  - On-premise/private

- **OS's**
  - Unix/Linux
  - Windows
  - MacOS

- **Programming Languages**
  - Python
  - R
  - Java
  - C/C++

- **Scripting Languages**
  - Unix Shell
  - awk
  - Perl
  - Julia

- **Math Computation Tools**
  - MATLAB
  - Octave
  - Mathematica

- **Visualization Tools**
  - Tableau
  - QlikView
  - TibCO Spotfire
  - SAS JMP

- **Python Tools & Libraries**
  - Jupyter Notebooks
  - TensorFlow
  - PyTorch
  - Hadoop/Hive/Pig
  - Spark/Mlib
  - Flume
  - scikit-learn

- **Data Mining Tools**
  - MS Excel Data Mining
  - MS SQL Server Mining
  - SAS Enterprise Miner
  - Oracle Data Mining
  - IBM Watson

- **Statistical Analysis Tools**
  - IBM SPSS Statistics
  - Oracle R Enterprise
  - Stan
  - Miniconda
  - Statistica

- **Business Intelligence Tools**
  - SAS
  - SAP Business Objects
  - Predictive Analytics
  - KNIME
  - IBM Cognos
  - Angoss

- **ML Productivity Tools**
  - RapidMiner
  - Cloudera
  - Amazon ML
  - Azure ML Studio
  - Orange
  - IBM SPSS Modeler
  - DataRobot
  - Salford Predictive Modeler

- **Databases**
  - Oracle SQL
  - Microsoft SQL Server
  - PostgreSQL
  - Other SQL
  - MongoDB
  - Amazon DynamoDB
  - Couchbase
  - Other NoSQL

---

**Application Experience Comparison**

<table>
<thead>
<tr>
<th>Functional Business Area</th>
<th>Automation</th>
<th>Natural Language Processing</th>
<th>Robotics/Industrial Control</th>
<th>Numeric Prediction</th>
<th>Recommendation</th>
<th>Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting &amp; finance</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Administration &amp; operations</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Customer service support</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Human resources</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Legal</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Logistics &amp; supply chain</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Marketing &amp; promotion</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Product/Service development</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Production/Manufacturing</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Research &amp; development</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Sales</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Other/unspecified</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
</tbody>
</table>

---

**Level of expertise**

Experience in role: 15

9/4/2018
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Invoicing &amp; payments</td>
<td>- Strategic planning</td>
<td>- Order &amp; billing support</td>
<td>- Staffing</td>
<td>- Contracting</td>
<td>- Sourcing &amp; procurement</td>
<td>- Brand creation &amp; awareness</td>
<td>- Feature ideation</td>
<td>- Production scheduling</td>
<td>- Research</td>
<td>- Lead conversion</td>
</tr>
<tr>
<td>- Payroll</td>
<td>- Facilities management</td>
<td>- Technical support</td>
<td>- Employee development</td>
<td>- Risk management</td>
<td>- Inventory management</td>
<td>- Customer understanding</td>
<td>- Product/Service definition</td>
<td>- Product creation</td>
<td>- Experimentation</td>
<td>- Relationship management</td>
</tr>
<tr>
<td>- Budgeting</td>
<td></td>
<td></td>
<td>- Compensation</td>
<td>- Legal defense</td>
<td>- Distribution &amp; shipping</td>
<td>- Lead generation</td>
<td>- Testing &amp; deployment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The table represents various departments and their corresponding functions. The table is structured to show the alignment between different departments and their responsibilities.*