Rise of the Machine (Learning)

Stephanie Wilson

ANZTB - 2017
@Stefowilso
DON’T BE SCARED
DATA ANALYTICS
MACHINE LEARNING
PREDICTIVE ANALYTICS
VR
AI
NLU
DATA MINING
AND MORE...
BIG DATA
NEURAL NETWORKS
CHATBOT
NLP
DEEP LEARNING
Artificial Intelligence
The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

Machine Learning
Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can change when exposed to new data.

Chatbot
A computer program designed to simulate conversation with human users.

Natural Language Processing
Natural language processing (NLP) is the ability of a computer program to understand human speech as it is spoken. NLP is a component of artificial intelligence (AI).
FAUX MACHINE LEARNING
MACHINE LEARNING
CHATBOTS
THE PAST & FUTURE
FAUX
ex machina

WHAT HAPPENS TO ME IF I FAIL YOUR TEST?
WHAT DO YOU MEAN, YOU’RE A SETTING?

WHAT DO YOU MEAN, YOU’RE AUTOMATED?
THE FIRST FORAY
TESTING THE WAF
Web Application Firewall

Checking for:
- Dodgy people (Source IP, Geolocation)
- Dodgy input/signatures (HTTP headers, Cookies)

go.xero.com
MODES
Simulation vs Active

- **Simulation** = Imperva WAF is monitoring traffic, and will give us risk-based alerts based on the rules set, and will let the traffic through

- **Active** = Imperva WAF will enforce rules and give us risk-based alerts. High risk alerts with high probability (i.e. hardly ever legit) are blocked.
MODES
Learning vs Active

- **Learning** = Imperva WAF is monitoring that URL and collecting information to understand normal patterns (e.g. parameter length used, methods used, etc). Traffic is let through.

- **Active/Not learning** = Imperva WAF creates rules based off the behavior it has learned and throws low, medium, or high alerts (i.e. blocks traffic) based off those rules.
IT’S TESTING JIM, BUT NOT AS WE KNOW IT
WHAT DID WE LEARN?

- Our team really disliked testing without a purpose
- Management didn’t really understand the terms exploratory and regression testing
- Our known testing techniques didn’t work
- It’s all about the data
MACHINE LEARNING
TYPES OF LEARNING

Supervised Learning
Creating functions from known data sets, also known as labeled training data
Task Driven learning using Classification and Regression

Unsupervised Learning
Training Data no longer used and any manual labels are removed
Machine is learning on its own using data driven clustering

Reinforcement Learning
Algorithm is learning to react to an environment on its own. There is a decision making process and reward system in place.
COULD WE DO MORE?
INVOICING WITH MACHINE LEARNING?

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Qty</th>
<th>Unit Price</th>
<th>Disc %</th>
<th>Account</th>
<th>Tax Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subtotal | 0.00
GST       | 0.00

TOTAL 0.00
SIMPLE RULES VS MACHINE LEARNING

Would simple rules be good enough?

- Last used account code
- Most frequently used
- Default sales code

What Models could we use?

- Multinomial Naive Bayes
- Logistic regression

Invoicing Simulator in Python

- Common way to compare methods
- F1 scores (1 is perfect, 0 is opposite)
- 2000 Organisations + 2 accounts
- 150 Invoices
- How much data do we need?
THE SET UP
TRAINING MODEL

Input Data

<table>
<thead>
<tr>
<th>Features</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Payee</td>
<td>Description</td>
<td>AccountCode</td>
<td></td>
</tr>
<tr>
<td>Shell</td>
<td>Petrol</td>
<td>449</td>
<td></td>
</tr>
<tr>
<td>Starbucks</td>
<td>Coffee</td>
<td>420</td>
<td></td>
</tr>
<tr>
<td>Shell</td>
<td>Diesel</td>
<td>449</td>
<td></td>
</tr>
<tr>
<td>Farm Table</td>
<td>Coffee</td>
<td>420</td>
<td></td>
</tr>
</tbody>
</table>

Examples

Transform

Train

Blank Classifier

Trained Model

Credit - Brad Stillwell

@Stefowilso - ANZTB 2017
RESULTS

Account prediction with different models

@Stefowilso - ANZTB 2017
WHERE DOES THIS LEAVE US?
WELLINGTON CoP MISSION STATEMENT

"We aim to deliver experiences that delight customers by influencing the Xero culture to be more quality oriented and transforming our software from 'good' to 'wow'!"
IT’S TESTING JIM, BUT NOT AS WE KNOW IT
THERE’S A LOT TO DO!

ANALYSIS
- What is a good candidate for machine learning?
- What does success look like? How do you measure it?
- What is a good data set for supervised learning?
- Do you need to A/B test?
- How do you move into production? Do you need a target customer type?

BE CURIOUS
- This is a young field, learn about it and what other people are doing
- How can this new technology help you to do your job?
- How can ML and AI be applied to new features or designs?
- Can you improve the UX by reducing exceptions or possibility of customer error?
IF YOU WANT TO LEARN MORE...
LOOK EVERYWHERE

@Stefwilso - ANZTB 2017
CHATBOTS

HOW CAN I HELP?
CHATBOT LOLZ

Facebook Conversation

**Justin**
10:12am
What can I do?

**Xero**
10:56am
Hi Justin! Good question... but I might need a bit more info before I can help :) -LG

**Justin**
11:00am
Oops, I thought you were a bot 😂. When is the xero bot launching?

@Stefowilso - ANZTB 2017
HELPFUL CHATBOTS
HAL9000: #headsup branch 35 tests have been triggered for XD-LowCarbResolutionTake2

WTF: dev-adamg merged their own PR into branch CI (627)
'revert "XD-Megapode release"' https://github.dev.xero.com/Xero/Xero/pull/627

ReganY on Aug 9
waltbot get directions "3 Market Ln, Wellington" "Hashigo Zake Wellington"

Waltbot on Aug 9
From: 3 Market Ln, Wellington, 6011, New Zealand
To: Wellington, New Zealand
Head northwest on Market Ln toward Cable St (74 m)
Turn right onto Cable St (0.1 km)
Turn right onto Taranaki St (signs for Brooklyn) (86 m)
Sharp right onto Jervois Quay (signs for Route 1/Route 2) (0.5 km)
Turn left onto Willeston St (0.2 km)
Turn left onto Willis St (4 m)

ReganY on Aug 9
wait. that isn't right.
INDUSTRIAL INTELLIGENCE REVOLUTION
HUMANS DO THINGS DIFFERENTLY
DON’T BE SCARED
Thank you!

@Stefowilso
Stephanie Wilson

ANZTB 2017