Machine Learning & AI in Transport and Logistics

Frank Salliau & Sven Verstrepen

Logistics Meets Innovation

Vlerick Brussels – Nov. 15th 2017
Sci-fi in 2002
Reality in 2017
PREDPOL®
The Predictive Policing Company
Sci-fi in 1984
Reality in 2017
What can I help you with?
What can I help you with?

“When is my wife's birthday”

Which ‘wife’?

“Siri I'm bleeding really bad can you call me an ambulance”

From now on, I’ll call you ‘An Ambulance’. OK?

Cancel

Yes
“Alexa, check my calendar”  "Alexa, what is machine learning?"

“Alexa, play Spotify”
What is it all about?
What is Machine Learning?

“The science of getting computers to act without being explicitly programmed” - Andrew Ng (Stanford/DeepLearning AI)
Challenge

Recognize dogs in images
What a human sees

DOG

NOT A DOG
<table>
<thead>
<tr>
<th>01101111</th>
<th>01101110</th>
<th>01101111</th>
<th>01101110</th>
</tr>
</thead>
<tbody>
<tr>
<td>01100101</td>
<td>01110011</td>
<td>01100101</td>
<td>01110011</td>
</tr>
<tr>
<td>00100000</td>
<td>01100001</td>
<td>00100000</td>
<td>01100001</td>
</tr>
<tr>
<td>01101110</td>
<td>01100100</td>
<td>01101110</td>
<td>01100100</td>
</tr>
<tr>
<td>00100000</td>
<td>01111010</td>
<td>00100000</td>
<td>01111010</td>
</tr>
<tr>
<td>01100101</td>
<td>01110010</td>
<td>01100101</td>
<td>01110010</td>
</tr>
<tr>
<td>01101111</td>
<td>01110011</td>
<td>01101111</td>
<td>01110011</td>
</tr>
</tbody>
</table>

???

???
What do we need?

TRAINING DATA
Labeled training set (dog/not dog) > 1000 images

Untrained Neural Network
Prediction phase

Unlabeled image

Trained Neural Network

87% DOG
13% NOT DOG
Why is this booming now?
What are the drivers?
(BIG) DATA
Byte: one grain of rice
Byte : one grain of rice
Kilobyte : cup of rice
Byte: one grain of rice
Kilobyte: cup of rice
Megabyte: 8 bags of rice
Byte: one grain of rice
Kilobyte: cup of rice
Megabyte: 8 bags of rice
Gigabyte: 3 Semi trucks
Byte : one grain of rice
Kilobyte : cup of rice
Megabyte : 8 bags of rice
Gigabyte : 3 Semi trucks
Terabyte : 2 Container Ships
Byte : one grain of rice
Kilobyte : cup of rice
Megabyte : 8 bags of rice
Gigabyte : 3 Semi trucks
Terabyte : 2 Container Ships
Petabyte : Blankets Manhattan
Byte : one grain of rice
Kilobyte : cup of rice
Megabyte : 8 bags of rice
Gigabyte : 3 Semi trucks
Terabyte : 2 Container Ships
Petabyte : Blankets Manhattan

**Exabyte : Blankets west coast states**
Byte : one grain of rice
Kilobyte : cup of rice
Megabyte : 8 bags of rice
Gigabyte : 3 Semi trucks
Terabyte : 2 Container Ships
Petabyte : Blankets Manhattan
Exabyte : Blankets west coast states

**Zettabyte : Fills the Pacific Ocean**
Byte : one grain of rice
Kilobyte : cup of rice
Megabyte : 8 bags of rice
Gigabyte : 3 Semi trucks
Terabyte : 2 Container Ships
Petabyte : Blankets Manhattan
Exabyte : Blankets west coast states
Zettabyte : Fills the Pacific Ocean

Yottabyte : A EARTH SIZE RICE BALL!
<table>
<thead>
<tr>
<th>Prefix</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte</td>
<td>one grain of rice</td>
</tr>
<tr>
<td>Kilobyte</td>
<td>cup of rice</td>
</tr>
<tr>
<td>Megabyte</td>
<td>8 bags of rice</td>
</tr>
<tr>
<td>Gigabyte</td>
<td>3 Semi trucks</td>
</tr>
<tr>
<td>Terabyte</td>
<td>2 Container Ships</td>
</tr>
<tr>
<td>Petabyte</td>
<td>Blankets Manhattan</td>
</tr>
<tr>
<td>Exabyte</td>
<td>Blankets west coast states</td>
</tr>
<tr>
<td>Zettabyte</td>
<td>Fills the Pacific Ocean</td>
</tr>
<tr>
<td>Yottabyte</td>
<td>A EARTH SIZE RICE BALL!</td>
</tr>
</tbody>
</table>

- **Hobbyist**
- **Desktop**
- **Internet**
- **Big Data**
- **The Future**
1 Yottabyte
1 Xenottabyte
1 Shilentnobyte
1 Domegemegrottebyte
1 Icosebyte
1 Monoicosebyte
Where does all this data come from?
The Power of the Crowd

The Norwich line steamboat train, from New-London for Boston, this morning ran off the track seven miles north of New-London.

Select all images below that match this one:

Type the two words:

Verify
BIG DATA IN PLANES: NEW P&W GTF ENGINE TELEMETRY TO GENERATE 10GB/S
classification

SVC Ensemble Classifiers

kernel approximation

KNeighbors Classifier

SGD Classifier

Naive Bayes

Text Data

Linear SVC

<100K samples

>50 samples

get more data

regression

ElasticNet Lasso

SVR(kernel='rbf')

Ensemble Regressors

predicting a category

do you have labeled data

<100K samples

few features should be important

<10K samples

predicting a quantity

predicting structure

<10K samples

just looking

<10K samples

Randomized PCA

Isomap Spectral Embedding

LLE

tough luck

dimensionality reduction
Supervised Learning
Regression

Price (in 1000s of dollars) vs. Size (feet²)

- Price: 220
- Size: 1250

Graph showing the relationship between price in 1000s of dollars and size in square feet.
Classification

- Iris Versicolor
- Iris Setosa
- Iris Virginica

Petal width [cm]

Petal length [cm]
Deep Learning
Google Deep Dream
(sometimes nightmare)
Unsupervised Learning
Reinforcement Learning
WATCH WHAT I CAN MAKE PAVLOV DO. AS SOON AS I DROOL, HE'LL SMILE AND WRITE IN HIS LITTLE BOOK.
Getting Started
CRISP Methodology

Exploratory

Iterative
GET ALL THE INFORMATION YOU CAN, WE'LL THINK OF A USE FOR IT LATER.
Quality of data: Garbage In Garbage Out
Multi-disciplinary team
Examples in Transport & Logistics
Watson Supply Chain

Discover the future of supply chains. Read the IDC Report: The Path to a Thinking Supply Chain

Download the report

Your supply chain impacts everything from the quality, delivery and costs of your business’s products, to customer service and profitability. Watson Supply Chain can elevate your existing systems to provide greater visibility, transparency and insight into supply chain data and processes—so you can better predict and mitigate the disruptions and risks that threaten your competitive advantage.
Ahlers Supply Network Innovation & Analytics (ASNIA)

Using ML to identify horizontal collaboration synergies between multiple shipper networks
Transmetrics: replacing budgets with prediction

Transmetrics big data cargo platform – a rigorous approach deriving benefits from current and future data

1. Data uptake cleansing and enrichment
2. Demand forecast modelling
3. AI optimisation
4. Execution Controlling

Traditional tactical management relies on budgets
Transmetrics: predictive tactical management

Budgeting process: “Flying on paper maps, experience and gut feeling”
Transmetrics: “High precision flying assisted by data and AI”
Data quality improvement – status of data after AUTOMATED / AI processing

Original situation
• Ambiguity in interpretation of orders and pieces records
• Missing size information at order and piece level
• Mix of measurement units (ldm, m³, kilos, pallets)
• Missing or unreliable capacity information for linehauls
• Other issues (e.g. senders with multiple name spellings)

Achieved situation
• Clear and correct identification of all shipments and pieces (including ignoring of non-piece lines)
• For each piece calculated full set of measurements (height, width, length, weight, volume) – *which enables 3D loading factors!*
• Complete set of measurements: loading meters, surface/pallets (m²), volume (m³)
• Built AI algorithm to estimate missing capacity information
• Grouped senders by AI; among others found a customer with 330+ different accounts/names

Preliminary data quality assessment after AUTOMATED / AI processing

Good to very good
Methodology / Behind the scenes

1. Preprocess data
   - cleanse and disaggregate pieces/volume/weight
   - Remove unnecessary rows
   - Ensure one row=one piece

2. Estimate volume
   - Predict missing volume data

3. Estimate dimensions
   - Mine text fields,
   - Apply industry standards and business rules

4. Predict missing linehaul capacity
   - use historical data to predict capacities
   - 95% accuracy

Approaches used
   - Combinatorial optimization
   - Quadratic optimization
   - Gradient boosting trees and other regressions
   - Natural language processing
   - Named identity disambiguation
   - Expert input / business rules

3D Loading Factor estimation
   - Loading meters (m)
   - Loading area, pallets (m²)
   - Volume (m³)
   - Expected 90% accuracy after calibration
3D loading visualization
Loading factor optimization results

Real situation (as observed on the warehouse floor)

Loading factor calculated by using the data

- **Loading meters**: 13.6 m (100%)
- **Floor**: 30.4 m² (91%)
- **Volume**: 29.8 m³ (34%)
Sven Verstrepen

Head of Supply Network Innovation & Analytics

sven.verstrepen@ahlers.com

http://www.linkedin.com/in/svenverstrepen
Frank Salliau
Independent Data Scientist & Machine Learning Expert

http://www.mentat-it.be

frank.salliau@mentat-it.be

https://be.linkedin.com/in/franksalliau