Deep Learning in Text Mining and Image Recognition

Dr. Andreas Nawroth
Head of Artificial Intelligence
Agenda for data enrichment and model creation using text mining and other unstructured data

1. Introduction to Analytics & Artificial Intelligence
2. Deep dive into Deep Learning approach
3. Our approach to industrialize Text Mining
4. Deep Learning in Image Recognition
5. Question & discussion
1 Introduction to Analytics & Artificial Intelligence
Munich Re actively shapes the transformation of the (re-)insurance industry

New trends & risks
- Digitisation
- Longevity
- Climate Change

…

MARKETS

NEW

Emerging markets

New products for emerging risks

ESTABLISHED

Traditional reinsurance business

Incremental innovations

PRODUCTS

Efficiently running the traditional book while continuously exploring new products and markets
Example: Ad-hoc risk identification and quantification
Example: Extraction of timeline of events
In recent years, there has been many applications of text mining.
Deep dive into Deep Learning approach
A good Deep Learning approach consists of several blocks: data input, neural network model, performance evaluation, and feedback loop.

1. Data
Words in sentences are converted to vectors.

2. Neural Network Model
Universal function approximation, the hierarchical layers mimic human brain.

3. Performance evaluation
Measure of how well the model approximates the task, e.g., cross-entropy, F1 scores, etc.

4. Feedback loop
Key to gain additional training data and improve model performance.

As compared with traditional Machine Learning model with relatively simpler representation, e.g., SVM, etc.; neural network model has deeper & more layers, bring more accurate results, and can deal with text with variable length. Training of the deep neural network is becoming easier nowadays due to improved algorithm and better computational powers.
Simple feedforward networks, feed info straight through nodes of networks

Feedforward Networks

\[
\begin{align*}
\text{inputs} & : x_0, x_1, \ldots, x_N \\
\text{target} & : y
\end{align*}
\]

\[
\begin{align*}
W_{ij} & \quad \text{Weights} \\
\sigma & \quad \text{Activation function} \\
\text{Neuron weighted output} & : z_{1k} = \sum_{j=0}^{N} W_{jk} \cdot x_j \\
\text{Neuron activation, e.g., sigmoid} & : a_{1k} = \sigma(z_{1k}) = \frac{1}{1 + e^{-z_{1k}}}
\end{align*}
\]

Simple Feedforward networks feed information straight through the nodes of the network, and never touching a given node twice. It has no notion of order in time, and the only input it considers is the current example it has been exposed to.
Deep Learning moves manual work complexity to model complexity

Deep learning offers a flexible framework to approximate complex & abstract tasks by automating the feature generation process.

Move manual work complexity to model complexity

<table>
<thead>
<tr>
<th>Classical programming</th>
<th>Machine Learning</th>
<th>Deep Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data preparation</td>
<td>Feature engineering</td>
<td>Rules Creation</td>
</tr>
<tr>
<td>manual</td>
<td>manual</td>
<td>automated</td>
</tr>
</tbody>
</table>

Universal approximation theorem\(^1\)

A feed-forward network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of \(\mathbb{R}^n\).

See tasks as functions

See complex tasks, involving structured or unstructured data as a function to be learned by the network.

Example: image models

Hierarchical abstraction\(^2\)

Neural network architectures typically process the data by adding complexity at each step of computation. This allows for modular re-use of models sharing similar core task.

1. Universal Approximation Using Feedforward Neural Networks: A Survey of Some Existing Methods, and Some New Results, Neural Networks, Scarselli et al., 1998
2. Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks, Honglak et al., ACM 2011
What is Text Mining: unlock the value of data in unstructured files

Unstructured data as input

Text Pre-processing
- OCR
- Tokenization
- Word Embeddings

Text Mining

Identify information
- Named Entities
- Relation
- Sentiment

Structured data as output

Deliver result
- Structured data in predefined format
- Risk Indices

Evolution of approaches

Look-up methods

Rule-based approach

“Traditional” Machine Learning approach

Deep Learning approach
One most used text mining techniques: Named Entity Recognition (NER), solves “Who exactly, When exactly, What exactly”? 

Illustrative example

Munich Re is expecting a loss of €1.4bn due to Hurricanes Harvey, Irma and Maria, for the third-quarter in 2017.

Desired outcome

Who: Munich Re

What: loss of €1.4bn

When: third-quarter in 2017

Data enrichment with additional data source (wiki, etc.)
NER challenge: how to extract the right info with the consideration of context?

With Munich Re’s MIRA Digital Suite, life insurers are massively reducing the effort and expense involved in applications and claims. CLARA, for example, halves the average time taken to settle disability claims.

Republican lawmakers still think Google is biased against conservatives, Google still claims that it’s not. The news agency reports.

https://www.healthinsuranceproviders.com/what-is-a-health-insurance-claim/

Fast innovations of text mining enable faster and more efficient info extraction for enhanced data quality and process automation

**Milestones of text mining**

- **2008**: NLP (almost) from scratch, Multi-task learning
- **2013**: Word embedding (faster)
- **2013**: Neural networks for NLP
- **2014**: Sequence-to-sequence framework
- **2015**: Attention
- **2015**: Memory-based networks
- **2018**: Pre-trained language models

**Enhanced efficiency**: With this framework, Google in year 2016 started to replace its monolithic phrase-based machine translation (MT) models with neural MT models (Wu et al., 2016), replacing 500,000 lines of phrase-based MT code with a 500-line neural network model.

**Reduced limitation**: Enables learning with significantly less data, only require unlabeled data.
TM framework captures recent AI innovations – text mining processes become more similar to human brain processes

Models evolved from rule-based to those represent human brains*

Deep learning neural network applied in text mining (illustrative)

Predicted label for words

Deep Learning model for Name Entity Recognition

Text

* Based on illustration on wired.com
Our approach to industrialize Text Mining
Scalable text mining platform: a flexible 3-tier platform, ensuring the optimal usage by internal and external business clients.

- Clients can run core components as standard functionalities.
- For client-specific needs, our text mining team is contracted to provide services (configuration and customization, incl. pipelines with multiple core components and data feeding).
- Clients can run PoC on platforms and evaluate performance and success of various approaches, including self-configured workflows.
Annotation tool enables flexible and easy way for information capturing

- Basic info
- Loss dates
- Loss location
- Risk info
- Parties involved
- Experts
- Circumstances of loss
- Legal issue
- Financials
Application example: we use text mining to detect organizational name and supply chain info.

Text mining service to detect organizational name and supply chain info

Documents → Customized service → Visualized supply chain info and risk → Underwriters use those info for business interruption risk for companies

Company Name extractor + Site locations + Transportation mechanism → Supply-receive relations (site to site relationship) + Risk Grading per site
Starting at a broad level... Energia provides a birds eye view allowing the user to select all the oil and gas infrastructure across the USA.
Data Lake Integration: text mining module extracts four types of info to enrich the existing documents

Info extraction

- Text Extractor
  - pdf
  - (embedded text)

- Upload UI
  - pdf
  - (embedded text)

- View PDF
  - pdf

- 4 Entity Extractors
  - Company
  - Location
  - Person
  - Date

- Table of Entities

- Keyword Index
  - Company
  - Location
  - Company & Location

- Feedback UI (check for false-positives)
  - Company
  - Location
  - (Person)
  - (Date)

Example: Find all PDF files in the Data Lake having the keywords “Munich” AND “BMW”
Demo: bring text mining to everyday tools such as Excel for everyone to access
Demo: Knowledge Graph to better understand organization profile
Our goal: enable access to insurance-specific and the most state-of-the-art text mining technology for both internal & external clients

Text mining at finger tips: everyone (data scientists or not) at Munich Re can run text mining for their everyday work, conduct experiments and PoCs with various AI approaches and models

Augmented underwriting: aided by text mining, underwriters analyze the text data from client on near real-time basis and provide immediate response to client requests

Client could access text mining modules through fast API

Packaged applications or customized use cases for clients
Deep Learning in Image Recognition
Catastrophe related challenges of a primary insurer

High volume of incoming claims

Difficulties in short term planning. Challenges in securing and placement of resources.

Restricted access to affected areas and limited ability to obtain ground truth

Maintaining Customer Service Levels – Keeping the Promise!
Our approach

- High-resolution aerial imagery
- Deep Learning
- Video Inspection Tool
- Simple web tool or API
Transfer Learning

Modern deep neural networks exhibit a curious phenomenon: when trained on images, they all tend to learn first-layer features that resemble either Gabor filters or color blobs. The appearance of these filters is so common that obtaining anything else on a natural image dataset causes suspicion of poorly chosen hyperparameters or a software bug.

Transfer learning refers to the situation where what has been learned in one setting (e.g., distribution $P_1$) is exploited to improve generalization in another setting (say, distribution $P_2$).

In transfer learning, the model must perform a different task, but we assume that many of the factors that explain the variations in $P_1$ are relevant to the variations that need to be captured for learning $P_2$. 

Demo: Deep learning approach applied for images
... allows a modern and tailored hurricane claims approach of an insurer.

No exterior damage, no building damage

Exterior damage

Light roof damage

Major damage to structural damage

No damage to roof, however fallen trees, broken fence or potential siding damage

Roof damage limited to shingles being ripped off

Roof damage can range from deeper layer damage with sheet wood visible, to roof structure severely damaged

Building fully not liveable anymore, typically resulting in total damage

One touch
Handle via remote inspection video app

Desk claim
Desk Adjusting by Inside Adjuster

Classic claim
On-site inspection and adjusting

Desk claim
Remote Adjusting by Inside Adjuster

New roof claims approaches possible
Value-add for insurers applying Remote Industries solution

- Higher claimant satisfaction
- Lower claims cycle time
- Faster settlement time / less “touches”
- Less field adjuster expenses (or more adjustments per day)
- Loss mitigation / better loss development
- Better overall catastrophe handling

Validated with 5 clients in 3 hurricanes in 2017/2018
Any application of AI within and outside of Munich Re shall be ethically justifiable and socially desirable. Therefore we adhere to the following principles:

- Human-centric AI
- Non-discrimination
- No Unacceptable Causality
- Data Governance for AI

We want to ensure that we satisfy these principles and maximize the benefits of AI while minimizing its risks. Thus we establish the following measures:

- AI Governance
- AI Validation
- Explainable AI
- Communication of AI

At the same time, we continue to comply with existing Munich Re guidelines as well as legal, social and cultural standards in all countries in which we operate.
Please feel free to contact me for further details!

Dr. Andreas Nawroth (ANawroth @ munichre.com)