Now and next in applied AI

Luka Crnkovic-Friis
CEO, Peltarion
About Peltarion

Founded in 2005, creators of an operational deep learning platform aimed at engineers and developers.

Small selection of customers.
What happens when the cost of deploying AI goes from $1m to $1k?
AI everywhere
supercharging every aspect of life
Our mission is to help that happen

One platform that exponentially lowers cost and skill requirement for solving real world problems with Deep Learning.

Allowing organisations to distribute AI access to domain experts and problems owners.
Why is AI so hard to get out of the lab?
The research view of AI
The operational view of AI

- Configuration
- Data Collection
- Data Verification
- Machine Resource Management
- Serving Infrastructure
- Monitoring
- Feature Extraction
- Model
- Analysis Tools
- Process Management Tools
Stopped trying at 01:12:36

Model deployed and tested at 00:07:55
The Original Story

- Early self-service user of the platform.
- Non-data scientists. (Medical engineers)
- Melanoma detection
- Worked, sort of, but not good enough

Atopic dermatitis and Eczema

- Measuring the skin barrier
- Predictive of eczema and other atopic conditions
- Can be applied to adults, but also babies at birth where preventive intervention is possible

Automated quality control

- Check of quality of the probe pins
- Fully automated now - video microscope & Peltarion platform
- Created by a summer intern
Current trends in AI
Unreasonable effectiveness of BERT

- Attention is all you need, 2017, Google
- Transformer: SOTA in machine translation
- BERT, encoder part of the Transformer
- 100+ million parameters
- Learns semantic embedding, the meaning of words
- Self-supervised learning without labelled data

Understands the contextual meaning in ways never possible before!
Paris Hilton stays with Orlando Bloom

Orlando stays at the Hilton in Paris
<table>
<thead>
<tr>
<th>Score</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9103</td>
<td>Will artificial intelligence really end the human race?</td>
</tr>
<tr>
<td>0.8362</td>
<td>Is technology destroying humanity?</td>
</tr>
<tr>
<td>0.8295</td>
<td>Is artificial intelligence really the end of humanity?</td>
</tr>
<tr>
<td>0.8108</td>
<td>Will the world end today?</td>
</tr>
<tr>
<td>0.8034</td>
<td>Will humans face extinction?</td>
</tr>
<tr>
<td>0.8021</td>
<td>When will the human race stop destroying earth?</td>
</tr>
<tr>
<td>0.7988</td>
<td>When will humanity end?</td>
</tr>
<tr>
<td>0.7980</td>
<td>Would losing net neutrality be the death of the Internet?</td>
</tr>
<tr>
<td>0.7976</td>
<td>Will the world end?</td>
</tr>
<tr>
<td>0.7976</td>
<td>Will the World end?</td>
</tr>
<tr>
<td>0.7970</td>
<td>Without net neutrality will the internet die?</td>
</tr>
<tr>
<td>0.7923</td>
<td>Will religion ever disappear?</td>
</tr>
<tr>
<td>0.7919</td>
<td>Is humanity destined to doom?</td>
</tr>
<tr>
<td>0.7914</td>
<td>How do you think human civilization will ultimately end?</td>
</tr>
<tr>
<td>0.7912</td>
<td>Will the universe end?</td>
</tr>
<tr>
<td>0.7911</td>
<td>How one can end the whole world?</td>
</tr>
<tr>
<td>0.7908</td>
<td>Will political correctness kill America?</td>
</tr>
<tr>
<td>0.7893</td>
<td>When will the human race end?</td>
</tr>
<tr>
<td>0.7886</td>
<td>Can the world end?</td>
</tr>
<tr>
<td>0.7876</td>
<td>Will terrorism in the world ever stop?</td>
</tr>
</tbody>
</table>
We’ve seen a lot of great progress in NLP recently:

BERT (Devlin et al., 2018)
RoBERTa (Liu et al., 2019)
XLNet (Yang et al., 2019)
ALBERT (Lan et al., 2019)
Reformer (Kitaev et al., 2020)

But it has been mostly limited to English.

What about Swedish, Norwegian, or Somali?
Cross-lingual transfer with XLM-R

Results on XNLI, a dataset for evaluating cross-lingual transfer, which in this case means training on English and evaluating on other languages.

<table>
<thead>
<tr>
<th>Model</th>
<th>D</th>
<th>#M</th>
<th>#lg</th>
<th>en</th>
<th>fr</th>
<th>es</th>
<th>de</th>
<th>el</th>
<th>bg</th>
<th>ru</th>
<th>tr</th>
<th>ar</th>
<th>vi</th>
<th>th</th>
<th>zh</th>
<th>hi</th>
<th>sw</th>
<th>ur</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>mBERT</td>
<td>Wiki</td>
<td>N</td>
<td>102</td>
<td>82.1</td>
<td>73.8</td>
<td>74.3</td>
<td>71.1</td>
<td>66.4</td>
<td>68.9</td>
<td>69.0</td>
<td>61.6</td>
<td>64.9</td>
<td>69.5</td>
<td>55.8</td>
<td>69.3</td>
<td>60.0</td>
<td>50.4</td>
<td>58.0</td>
<td>66.3</td>
</tr>
<tr>
<td>XLM (MLM+TLM)</td>
<td>Wiki+MT</td>
<td>N</td>
<td>15</td>
<td>85.0</td>
<td>78.7</td>
<td>78.9</td>
<td>77.8</td>
<td>76.6</td>
<td>77.4</td>
<td>75.3</td>
<td>72.5</td>
<td>73.1</td>
<td>76.1</td>
<td>73.2</td>
<td>76.5</td>
<td>69.6</td>
<td>68.4</td>
<td>67.3</td>
<td>75.1</td>
</tr>
<tr>
<td>XLM-R</td>
<td>CC</td>
<td>1</td>
<td>100</td>
<td>88.8</td>
<td>83.6</td>
<td>84.2</td>
<td>82.7</td>
<td>82.3</td>
<td>83.1</td>
<td>80.1</td>
<td>79.0</td>
<td>78.8</td>
<td>79.7</td>
<td>78.6</td>
<td>80.2</td>
<td>75.8</td>
<td>72.0</td>
<td>71.7</td>
<td>80.1</td>
</tr>
</tbody>
</table>

On our data, we got 85.7% accuracy training and evaluating on Swedish. With cross-lingual transfer from English to Swedish, we got 84.4% accuracy.

We can achieve good performance even without labeled data in your language.
A peek inside XLM-R

- Sample five sentences from the XLNI dataset
- Compute XLM-R representations and project to 2D with UMAP

I didn’t know what I was going for or anything, so was to report to a designated place in Washington.

Я не знал, что мне предстояло сделать и все такое, так что я должен был сообщить в указанное место в Вашингтоне.

他说，妈妈，我回来了
และเขาพูดว่า，มам่า ผมอยู่บ้าน
And he said, Mama, I’m home.
Use case: Political party classifier

HMM... NO.
All documents from Swedish parliament is publicly available
- Motions, for instance
- We collected all motions from 2010 onward
- 32 153 motions
Teaching AI politics

What could possibly go wrong?

Author identification
- “Which party wrote this motion?”
- Classification problem with eight classes
- Swedish corpus
- Used XLM-R

Minor pre-processing
- Replace party names with “Partiet”
- Split motions into chunks of 1-4 sentences
Did the model learn politics?
Swedish party classifier

Which Swedish party wrote your text? We trained a multilingual classifier on publicly available motions from the Swedish Riksdag. While the model is trained on Swedish data, the underlying model, XLM-R, handles 100 languages, so try whatever language! To learn more, check out our blog post, and if you're more interested in the technical details, read our deep dive on multilingual models.

Write your political text here!

Note: The model may not be as reliable on text that is significantly (semantically or syntactically) different from the political texts it was trained on. Regardless, it's fun to try!

https://swepol.research.peltarion.com/
The Future
Explainable predictions

Improve **model interpretability** and build **user trust**

- How to help data scientists to understand why some predictions are wrong?
- How to visualize explanations for domain experts in an intuitive way for text?
- Promising solution: WT5 model from Google - ask the neural net
One model to rule them all

Zero-shot learning. No additional training necessary.

- Can solve multiple tasks, without additional training
- The user shows the model what it expects and it has an internal “reasoning” process to try to answer as expected

Promising solution: **GPT-3** from OpenAI

- It has more than twice the number of weights than the number of neurons in the human brain.
- It moves closer to a single model for generic tasks, including common-sense reasoning, arithmetic, and translation.
- The ability for zero/few-shot learning is moving much closer to how humans think.
Summary

The Promising Part

- Big models are awesome, first step to AGI & no end in sight (stack more layers). Free.
- Modern platforms are allowing domain experts to apply modern AI techniques, and make a difference.

The Worrying Part

- OpenAI’s GPT-3 model costs about $1M to train.
- All the data you can eat => implicit bias.