Automaise In Short

Automaise provides an AI+NLP based Automation Platform to deploy AI virtual employees capable of automating several business processes and tasks, both customer facing and backoffice.

Abstract

Neural nets usage is all the rage, however several challenges arise once we apply it to a real world problem:

- Enabling simplified supervision

- Making neural nets keep learning by themselves

- Unbalanced training datasets

- Guarantee the quality of the training dataset

Supervision by Common People

- Use common sense terminology

- Make the most out of the operator/supervisor work

- Infer every time possible

Enabling simplified supervision

Sentence	Typification	Completion	Remove
wad asd asd	11	OK	✕
enbtão não quero nada	4	OK	✕
então não quero nada	4	OK	✕
oi bons dias!	1	OK	✕
pt50000700000004	11	OK	✕
Moveis, eletrodomesticos	11	OK	✕
Posso saber quem és tu?	9	OK	✕

Current Model

Typification Identifier	Typification Description
1:	Hello
2:	NoAction
3:	Yes
4:	No
5:	GoBack
6:	Thanks
7:	Goodbye
8:	Operator
9:	AboutAnna
10:	NotSureDontKonw
11:	NotApplicable

+

Automaise © 2018

Making neural nets keep learning by themselves

Everyone says that by making use of AI/Neural nets their solutions will keep learning from user/customer interactions, but the truth is that few or no company complies with this commitment.

The Virtuous Learning Cycle
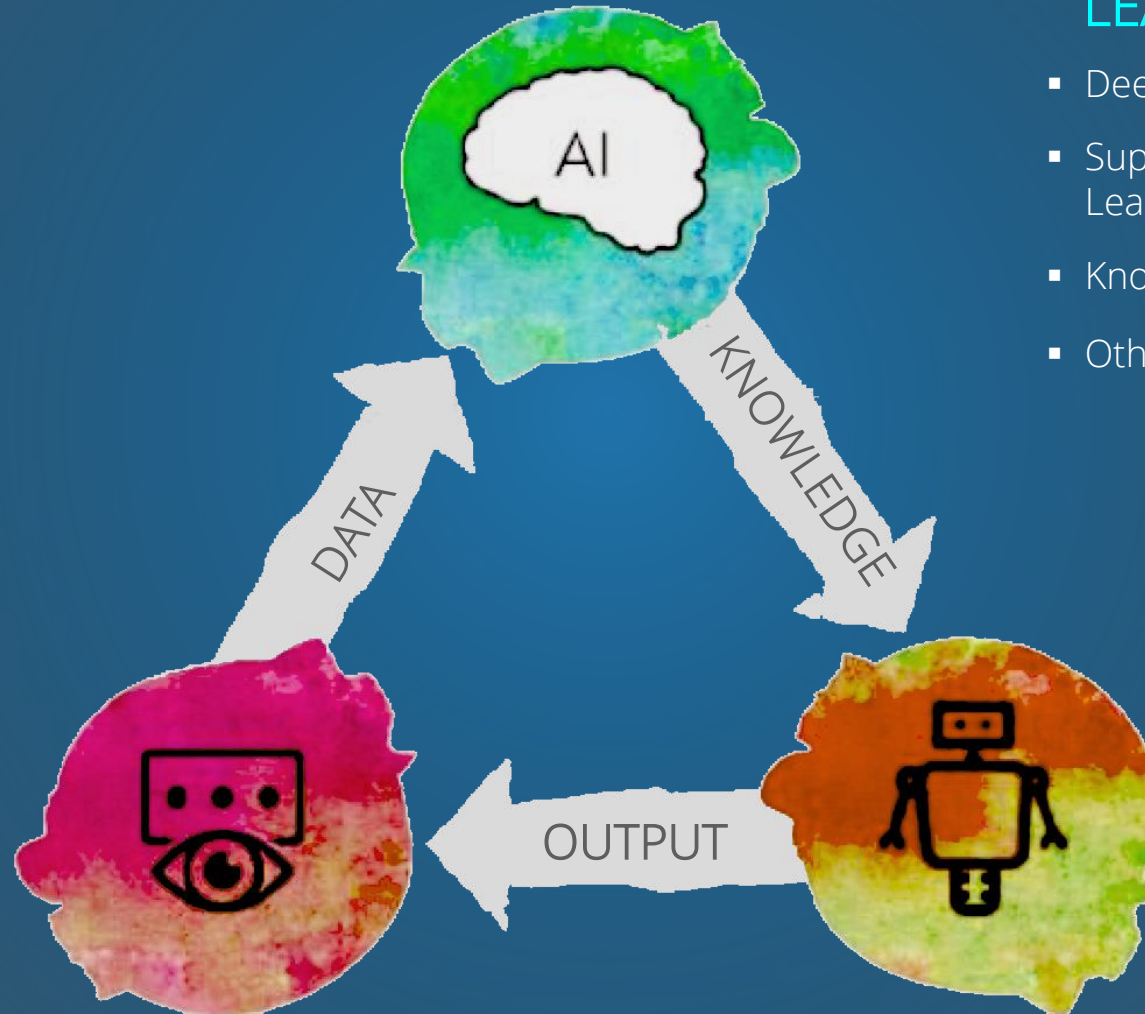
LEARN
- Deep Learning and Other AI Technics
- Supervised Learning + Reinforcement Learning + Self Taught
- Knowledge Base
- Other Rules

OBSERVE
- Structed Text Data
- Unstructured Data
- Computer Vision – Documents OCR
- Train Data Models

EXECUTE
- AI Process Automation
- Automation Flow AI
- APIs and Other Integrations
- Tools and Routines

KNOWLEDGE

DATA

OUTPUT

AI

automaise
YOUR BUSINESS AUTOMATED

DSPT
DATA SCIENCE PORTUGAL

Unsupervised Learning

Techniques / Approaches

- Clustering (K-means, EM Clustering, Affinity Propagation, etc.)

- Neural Networks (Autoencoders, Deep Belief Nets, etc.)

- Automaise's proprietary approach: The "Teaching Committee"

The "Teaching Committee"

- Convolutional, Recurrent (+LSTM), Hierarchical Attention Network, SVM, fastText, EntityNet, etc.
- Heuristic based on several parameters:
  - Consensus
  - Confidence
  - Model weights
  - Etc...

Unbalanced training datasets

Unbalanced that set pose a problem on neural nets learning. On almost every business certain subjects happen with a (much) higher frequency than others, therefore the training datasets become unbalanced between classes, which will make neural nets biased.

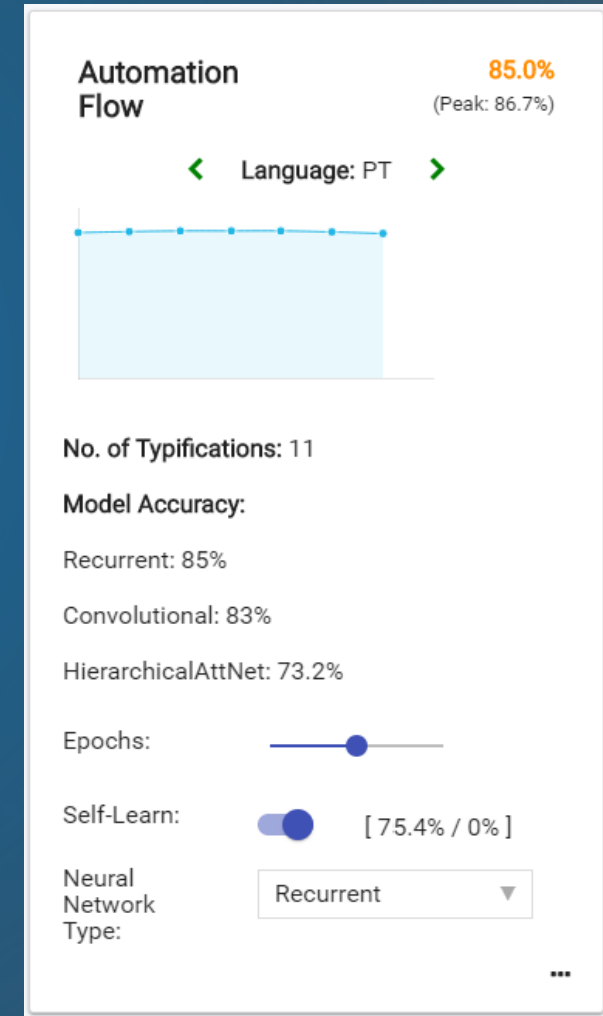# Imbalanced Data

Techniques / Approaches:

- Resample differently

- Try different metrics / classes

- Merge less frequent classes into a "other matters" class

- Use penalized models

- Automaise proprietary approach... ☺

Guarantee the quality of the training dataset

As training data sets are mostly supervised by humans, sometimes without a strong commitment with the project/task, classification errors happen quite frequently which end up damaging neural nets accuracy.

Train Dataset Quality Assurance

- Performance dashboard

- Mind peak performance

- Detect classification conflicts

- Automatically detect potential misclassified train samples

automaise
YOUR BUSINESS AUTOMATED

DSPT
DATA SCIENCE PORTUGAL

Automation Flow   **85.0%**
                        (Peak: 86.7%)

‹   Language: PT   ›

No. of Typifications: 11

Model Accuracy:

Recurrent: 85%

Convolutional: 83%

HierarchicalAttNet: 73.2%

Epochs:

Self-Learn: [75.4% / 0%]

Neural
Network   Recurrent   ▾
Type:

...

Some Useful Pointers

- https://github.com/fendouai/Awesome-Text-Classification

- https://github.com/brightmart/text_classification

- http://www.wildml.com/

- https://spacy.io/

Thanks!

Ernesto Pedrosa • +351 968690232 • ernesto@automaise.com