WAYS OF USING ARTIFICIAL INTELLIGENCE IN THE BANK DOMAIN

Zyzygy Inc

PREPARED BY
Edgardo Promenzio
Ana Maguitman
Axel Soto

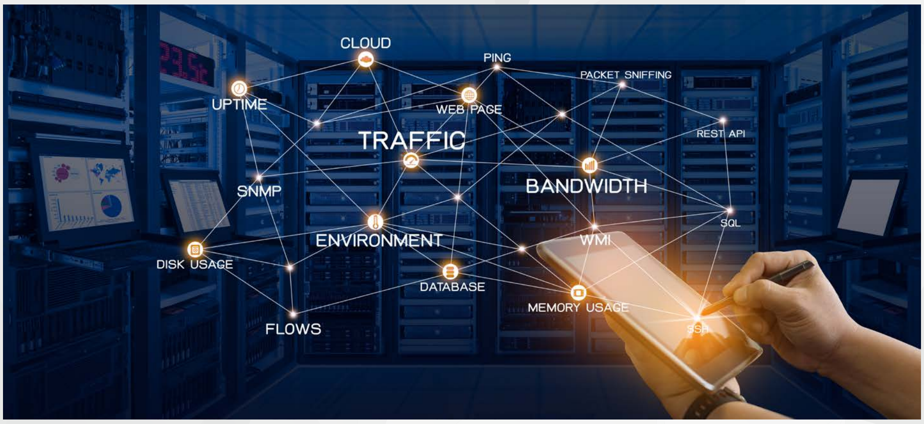
Ways of Using Artificial Intelligence in the Bank Domain

Artificial Intelligence (AI) represents a major game-changer compared to the ways the traditional bank industry has been operating on. Possibilities offered by AI have an impact across the entire organization by reducing human experts' time in repetitive tasks, which yields improved resource management, improved customer experience and more efficient processes throughout the entire organization. Based on a real use case with a major bank organization, we describe here major tasks, challenges and opportunities.

Feature engineering

Machine learning algorithms take a set of features to make predictions [Domingos 2012]. Feature selection techniques help to interpret the data and reduce the effects of the constant dimensional growth of the dataset [Chandrashekar and Sahin 2014]. Some feature selection techniques are based on determining the existence of dependencies between the potential features and the labels to be predicted. This allows to identify and filter irrelevant attributes. Among the best known filtering techniques are those based on correlation criteria [Jiang et al. 2018] and those that use the notion of mutual information [Sharmin et al. 2019]. The main disadvantage of filtering methods is that the selected subset of attributes could have many redundant attributes. As a consequence, filtering methods can be integrated with "wrapper methods", which are based on evaluating the effectiveness of different subsets of features in predicting the labels [Wah et al. 2018]. The main disadvantage of these methods is their computational time. As an alternative, embedded methods incorporate the selection of attributes into the training process itself [Liu et al. 2019].

In addition to applying feature selection techniques to remove noisy attributes, it is important to generate new attributes that are not explicit in the dataset. Some feature extraction techniques that are typically applied to generate new attributes include traditional approaches such as those based on principal component analysis [Schölkopf et al 1997] and more recent ones, such as word embeddings [Mikolov et al. 2013, Devlin et al. 2019].

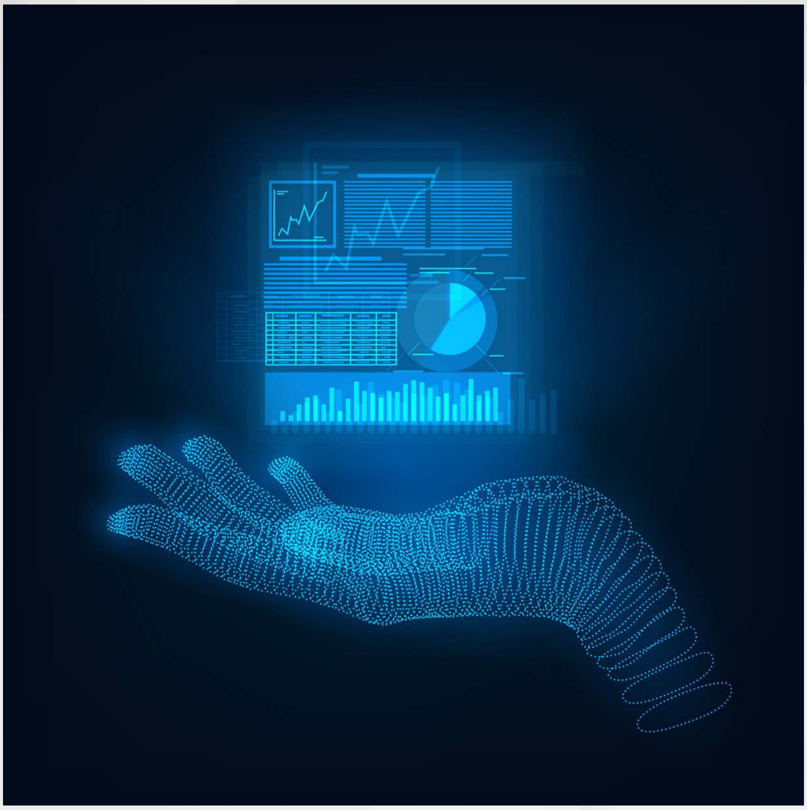


Information extraction

The main goal of information extraction is to identify useful information, such as entities or relations, in unstructured data with the purpose of transforming it into structured data [Grishman, 2015]. The task of named entity recognition (NER) is to find mentions of named entities in a text. Possible named entities that may be useful in the bank domain include organizations, codes, dates, money, people, and places. In some cases, the entities have very specific grammars (such as for certain banking system codes) that might make it possible to recognize them independently of the given context. Although it is possible to write rules by hand to define patterns for NER, this is quite difficult because these rules should take into account both the specific tokens and context. Moreover, such rules require continuous maintenance and curation so that they reflect current terminologies.

The state of the art in NER relies on the use of a feature-based sequence classifier such as a Conditional Random Field (CRF) or a neural sequence classifier such as a bi-LSTM model. These classifiers are trained to label the tokens in a text as outside or inside any entity.

Another important task is the identification of relations that exist among the detected entities. The relations can take a general form or be specific to the bank domain. The use of domain-specific ontologies can aid the process of domain-specific relation extraction.



Classification

The application of machine learning techniques in a bank help desk system can lead to increased productivity, better quality of service and improved customer satisfaction. In particular, automatic classification and routing of tickets to the responsible agents can significantly reduce ticket resolution time.

Associating a ticket with an agent or service is a multiclass document classification problem and can be addressed using supervised machine learning algorithms such as neural networks, support vector machine, decision trees, etc. These algorithms require access to training data in the form of labeled ticket information.

Similarly, associating regulations from the central bank to subsystems that need to modify their behavior can also be tackled as a multiclass document classification problem.

Challenges and Opportunities

Challenges and Opportunities

While information extraction techniques and machine learning offer a wide-range of applications for the automatic processing of text, specific domains such as banking or financial ones may require additional domain knowledge to augment the effectiveness and allow intelligent decision making processes.

Some of the solutions that can help address these challenges include:

• **Ontological knowledge**: Domain knowledge can be usefully incorporated into the text mining and machine learning processes. Ontologies offer a useful vehicle to guide these processes by providing a specification of concepts and their relations in a specific domain. The encoded formal semantics in ontologies allows sharing and reusing knowledge and data.

• **Semantic annotation:** Semantic annotation is important in realizing semantic text mining and machine learning by bringing formal semantics to data. Annotated data is useful for these processes because data acquires a more formal and structured format through ontological terms and their relations. A typical use can be to annotate the classification labels. Clustering can also benefit from ontology-based annotations by enriching term vectors with ontological concepts and incorporating ontology semantics into similarity measures.

* **Semantic text mining**: An ontology-based approach for semantic text mining or machine learning has the purpose of using formal ontologies in these processes. This can be achieved by taking advantage of the concepts and relations encoded in the ontology to constrain these processes, by making them more precise, and to semantically augment them, by making them more complete. For instance, in a classification task, an ontology can specify the consistency relations or help relate different terms into a common concept. By constraining the search space and identifying semantic relations, classification can be more accurate. Another example is ontology-based information extraction, where the extracted information is a set of annotated terms from the documents with the relations defined in the ontology. This allows to organize the extracted information in a formal and structured way using an ontology representation.

References

Chandrashekar, G., Sahin, F. (2014) A survey on feature selection methods. Computers & Electrical Engineering 40.1:16-28.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT.

Domingos P. (2012) A few useful things to know about machine learning. Communications of the ACM. Oct 1;55(10): 78-87.

Grishman, R. (2015). Information extraction. IEEE Intelligent Systems, 30(5), 8-15.

Jiang, Liangxiao, et al. (2018) A correlation-based feature weighting filter for naive bayes. IEEE transactions on knowledge and data engineering 31.2: 201-213.

Liu, Haoyue, MengChu Zhou, and Qing Liu (2019) An embedded feature selection method for imbalanced data classification. IEEE/CAA Journal of Automatica Sinica 6.3:703-715.

Mikolov, Tomáš, Wen-tau Yih, and Geoffrey Zweig (2013) Linguistic regularities in continuous space word representations. Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics: Human language technologies.

Schölkopf, Bernhard, Alexander Smola, and Klaus-Robert Müller (1997) Kernel principal component analysis." International conference on artificial neural networks. Springer, Berlin, Heidelberg.

Sharmin, Sadia, et al. (2019) Simultaneous feature selection and discretization based on mutual information." Pattern Recognition 91:162-174.

Wah, Yap Bee, et al. (2018) Feature Selection Methods: Case of Filter and Wrapper Approaches for Maximising Classification Accuracy." Pertanika Journal of Science & Technology 26.1.