Annals of "Dunarea de Jos" University of Galati
Fascicle I. Economics and Applied Informatics
Years XXVI – nº2/2020
ISSN-L 1584-0409 ISSN-Online 2344-441X
www.eia.feaa.ugal.ro

DOI https://doi.org/10.35219/eai15840409110

Empirical Insights on Cloud Services for Machine Learning Applications

Adrian MICU*, Marius GERU, Angela-Eliza MICU***, Alexandru CAPATINA****, Constantin AVRAM*****, Robert RUSU******

ARTICLE INFO

Article history:
Accepted August 2020
Available online August 2020
JEL Classification
L81, M31, C80

Keywords:
Machine Learning, Artificial Intelligence, Cloud Services, Computer Vision

ABSTRACT

As the volume of data increases, becomes more complex and valuable, people's limited capabilities present real challenges in deciphering and interpreting an increasingly unpredictable economic environment. In essence, Machine Learning is the artifact of artificial intelligence generated and shared mainly by the technological environment, where almost any information can be documented, measured and stored digitally, thus becoming data that can be processed to generate actionable information reusable in multiple spheres. of activity. The aim of this research is a comparative analysis of the main cloud services available for Machine Learning algorithms. The research results offer a dynamic vision to the researchers involved in the FutureWeb project, who are looking for the most efficient cloud platforms for the services offered by the AI Media platform.

1. Introduction

Machine learning refers to the techniques needed to work intelligently with a large amount of data, by developing useful algorithms to draw some conclusions about this data.

It would be impossible for a human user to perform the activities involved in Google, to analyze what we like and to redirect the searches to the useful links to us. This is where Machine Learning comes in, an integral part of AI artificial intelligence.

Artificial intelligence is a concept that refers to the ability of a system to learn on its own to adapt, to react in completely new situations, i.e. to act in an "intelligent" way. Applications with machine learning are the version of artificial intelligence currently applicable, characterized by the use of mathematical algorithms, which endows them with the ability to operate with a large volume of data, to learn on their own and thus be able to improve their actions. The key difference from pure artificial intelligence, which should act completely independently, is that Machine Learning-based applications operate under human control and within established parameters.

2. Theoretical background

Machine Learning is a concept introduced by Alan Turing, who also created the famous test of the same name. Based on the game of imitation, the test aims to answer the question of whether an application can behave like a thinking person (Turing, 2004).

Artificial intelligence is an extremely vast and complex field, including many subdomains, which often generates confusion related to the differences between AI, Machine Learning, Deep Learning.

Artificial Intelligence translates into sets of rules and the analysis of possibilities, the basis from which decisions are made in a system. Machine learning is a subdomain of artificial intelligence, where one learns from previous experiences, represented by data that will be combined with algorithms, to reach the final results. These final results refer to the construction of a model, most often specialized in carrying out an activity.

In the field of computer science, machine learning, which is a subset of artificial intelligence (Jordan and Mitchell, 2015), has been widely applied in areas such as natural language processing (NLP) (Guimaraes et al. 2018), speech recognition (Agarwalla and Sarma, 2016) and computerized visualization. Conventional machine learning approaches are limited in their ability to process data in their raw form (LeCun et al., 2015). The inability to process data is due to the fact that a considerable amount of knowledge in programming and

*, ****, *****, ****** Dunarea de Jos University of Galati, Romania, **Transilvania University of Brasov, Romania, ***Ovidius University of Constanta, Romania. E-mail address: marius@thecon.ro (M. Geru – Corresponding author)

fields is required for the design of a feature extractor (LeCun et al., 2015). Price et al. (2018) managed to create a nomenclature that automatically extracts and indexes their specific products and attributes.

Deep learning is a representation method that can be used to automatically extract sophisticated features at high levels of abstraction. This method can also learn from data with multiple levels of end-to-end representations (LeCun et al., 2015). By combining deep learning methods (neural networks) with computer visualization, specific elements or products from images can be extracted and used for managerial decisions or marketing strategies. A particular type of deep learning method that has been widely used is CNN (Convolutional Neural Network), which has managed to outperform other neural networks in terms of image classification (Krizhevsky, 2018), object detection and segmentation. them. Deep learning has allowed the development of applications based on computer visualization, for example, in the case of autonomous vehicles and in the automatic diagnosis of breast or skin cancer (Mehta et al., 2020).

The concept of Deep Learning was developed as an extension to the concept of Machine Learning, being based on the introduction of a large amount of data in a system of automatic processing of data and information, through neural networks. This technology can analyze the information in more detail. However, Deep Learning has a defining characteristic, namely the ability to remember past actions and to learn from one's own behavior. Among the best known Deep Learning applications are those that refer to the navigation system of self-driving cars or the ability of computers to recolor black and white images and videos, after learning what certain objects look like and how they are perceived by people (Maqueda et al., 2018). Another relevant example would be Google DeepMind, an intelligent system that uses algorithms based on Deep Neural Networks.

Therefore, Deep Learning is subordinated to Machine Learning, which in turn is part of everything that means Artificial Intelligence. The image below illustrates the relationships between them:

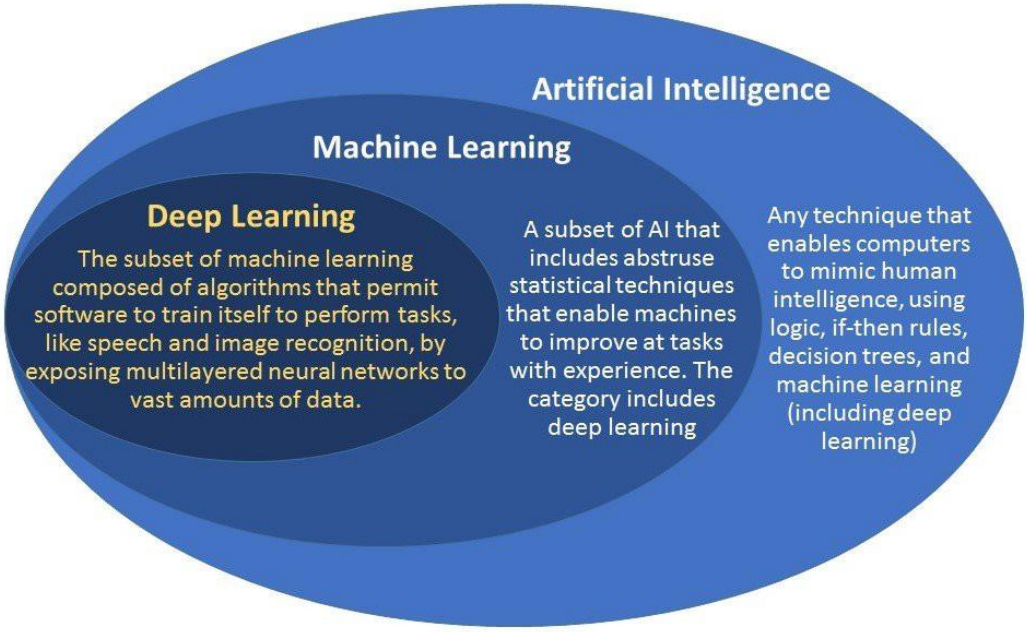

Figure 1. Positioning of Machine Learning
Source: original contribution

Machine Learning is used to teach a robot the difference between different items and then recognize them in pictures (Ciregan et al., 2012). Data Mining (or data extraction) will be used when a program is instructed to collect images of dogs and images of cats from the Internet and without necessarily using or understanding them. In the selection process count a number of parameters such as image quality, resolution, colors, an option could be to standardize and change the format to a black and white or change the images to the same size, identify duplicates and / or remove them.

There are a few common steps in any Machine Learning activity (Hodnett and Wiley, 2018):

• Input data - in any Machine Learning activity, the input data is vital and can be in the format: text, csv, google docs, etc. Understanding the data we have is the first big step towards a correct result, whether we are talking about the classification of images, text, sound, etc. When the results are not as expected, a common approach is to go back to the input data, because they usually have the answers. For example, we get to a point when the model built does not have the desired performance and has a very poor accuracy. This may mean that the input data is not sufficient. Or, simply, the data were not annotated properly.

• The sources from which the data are selected - must be as varied as possible, and also the volume of data is very important. Depending on the activity to be performed, the data set may need to be larger and larger in order to achieve satisfactory results.

• Data preparation - involves determining the quality of the data collected and most often, leads to the identification of missing data, or those data with properties far beyond the usual for the category to which they belong.

• Data abstraction - data will be abstracted by representation in a structured format, by the chosen algorithm. At this stage, elementary learning is achieved. The format we give to the data must be as simple as possible to process and understand through the training phases, in order to successfully achieve the purpose of classification. For example, if, based on characteristics such as age, sex, annual income, criminal record, it is necessary to determine whether someone is eligible to receive a loan, it is necessary that some data be in numerical format in order to make calculations. on them.

• Model training - involves choosing an appropriate algorithm, depending on the problem to be solved. To this is added the representation of the data in a form corresponding to the model. In this sense, the data are divided into subsets: training and test. The proportion in which the data is divided depends on how large the data set is, and other factors related to the algorithm, or the observations made in the previous phase on the data set. Training data is used to develop the model, as, based on it, it learns how to distinguish various categories. Data from the test set are used as a reference to measure the efficiency of the built model.

• Generalization - is the stage that generates conclusions after learning from the previous step. The evaluation of the model is an integral part of the generalization stage, which measures how accurately the model works on data that it has not used until then. Another way to determine the generalizing power of the model can be by testing it on data from a category that simply was not used at all in training. Improving performance is also part of the generalization stage, being the activity that ends by obtaining conclusions.

Depending on the nature of the data we have available and the nature of the problem to be solved using Machine Learning, we will encounter the following types of learning (Ayodele, 2010):

1. **Supervised** - Models for this type of learning are predictive. In other words, we give them from the beginning all the indications about what they have to learn. That is, they will have data labeled by a human user and will know, for example, in the classification of images, what is "a logo" and what is "a banner". All that remains is to learn distinctive features for each, to answer the question: "Why would he consider the one who made the annotation of the data that in a certain image is a logo and a brand?". This process involves identifying specific traits or attributes. In the case of images, these attributes are pixels, pixel formations; in the case of text classification, it can be words and word combinations.

2. **Unsupervised** - Training data is not labeled to determine what results are desired. We also look at models and how they work, but also what their purpose is to cluster data on similar attributes. As there is no desired result, there are no more important attributes than others. Unsupervised learning is used in the retail area to determine which product categories customers tend to buy more often.

3. **Enhanced** - It is the most ambitious type of learning and bears this name because there is a possible reward following sequences of actions. In the case of this type of learning, decisions are based on certain activity-specific requirements. Applications that use enhanced learning have a single purpose, that of making decisions to maximize a utility. There are applications that operate in a dynamic environment and that train constantly. As more knowledge is gained and more experience is gained, they become more capable of solving problems in the field. Enhanced learning uses chess applications. Here the involvement, for data annotation or guidance, from the human user is minimal.

3. Comparative analysis of cloud solutions for Machine Learning

Large software vendors see machine learning as a huge market for profit, which is why there are many vendors of machine learning tools and platforms, such as SAS and IBM, Oracle or Amazon for cloud solutions. From a business perspective, these tools can help solve problems, but asking the right questions, gathering the right assumptions, and measuring success is, in most cases, the role and focus of machine learning stakeholders with a mindset of business. The value of machine learning is especially appreciated by businesspeople based on proven contributions or achievements.

The applications that use Machine Learning are more and more diverse and cover many areas such as:

• Search engine optimization - Index pages on search engines according to the browser's preferences and profile. The constant evolution of search engines is aimed at improving the user experience - annually the algorithms are modified at least 500 times for maximum performance and to optimally satisfy searches. Google acknowledged years ago that it had more than 200 ranking factors, which played a vital role in how it ranks sites, and in the meantime things have only gotten more complicated. This is the main reason why search engine optimization techniques are constantly changing. RankBrain, the algorithm that represents a trend between 2017-2018, introduced by Google in 2015, takes quick steps in understanding human intentions. Its ability to learn gives it the ability to rank new keywords and group them by domain, and this will certainly influence search engine optimization in the future. Many of the searches that users do on a daily basis are new to the search engine, so Google has been able to develop an algorithm to link different concepts. Based on the fact that the new key phrases seem to resemble something he has seen in the past, he assumes or intuits the search results.

• Sales - quick identification of the products that are sold more frequently than the less popular ones, this can contribute to making the right decisions about the stored products and the size of the stock.

• Finance and banking - based on standardized answers, you can build a model that predicts which customers are more likely to pay their debts and bills.

• Medicine - diagnosis of cancer or other difficult-to-cure diseases, based on patients' symptoms.

Google Cloud Automachine Learning Vision

Google launched a new AI tool in 2017, called Cloud Automachine Learning Vision, designed to allow anyone to train their Machine learning systems on a photo database. To be functional, the user needs a multitude of tagged images, and up to 1000 units per month the service is available for free, but the software can support millions of units.

Cloud Vision offers two pre-trained models via an API and the ability to build custom models using Automachine learning Vision to provide case-specific flexibility. The Cloud Vision API allows developers to understand the content of an image by encapsulating powerful machine learning models in a REST API (Application Programming Interface), which is a set of definitions of sub-programs, protocols and tools for creating application architecture. REST type, resource oriented) easy to use. It quickly classifies images into thousands of categories, detects individual objects and faces in images, and reads the printed words contained in images. The user can create metadata in his image catalog, moderately offensive content or can allow new marketing scenarios by analyzing the feeling.

Automachine learning Vision allows developers with limited experience in machine learning to prepare high quality custom models. After uploading and labeling the images, Automachine learning Vision will prepare a model that can be adapted to user requirements. It offers increased model accuracy and faster time to create a production-ready model.

An important activity in image recognition is given by the identification of models. Once Google AI believes it has a good understanding of the links between the images the user has uploaded, it can be used to search for that model in new images, removing images that do not match the model. The ability to recognize models at the scale of a 10,000-entry database has immense interdisciplinary value. Oncologists have trained machine learning systems with images of breast cancer cells so they can detect the disease earlier. Neurologists have used algorithms to scan MRI to predict language development in children. Researchers at Stanford applied similar software to predict voting patterns in cities by correlating census data with the frequency of certain car brands.

Automachine learning Vision has increased the number of these projects, because early detection has huge potential in saving many lives, being able to discover new models and unproven correlations. With Automachine learning Vision, the limitations are primarily related to data collection, respectively the capture and correct marking of thousands of images for training. There are more ways to capture images than ever before: through mobile phones, drones, live feeds or social media, but with their development, there have been legislative inconveniences related to the right to capture certain images.

Amazon Recognition

Amazon Rekognition not only analyzes still images, but can provide insight into the videos. Amazon Rekognition provides users with a tool for continuous image recognition, being intuitive, easy to integrate and use, capable of:

• Object, scene, and activity detection - makes it easy to add images and video analysis to existing applications. By providing only an image or video to the Rekognition API, the service can identify objects, people, text, scenes, and activities.

• Facial recognition - identifies a person in a photo or video.

• Facial analysis - provides a very precise facial analysis and facial recognition. Faces can be detected, analyzed and compared for a wide range of use cases, including user verification, cataloging, person counting and public safety. (smile, open eyes, glasses, beard, sex).

• Identifying patterns - for example, the movement of athletes during a game for post-game analysis.

• Detect inappropriate content - identify inappropriate content in images and videos. Developers can use the returned metadata to filter out inappropriate content based on their business needs. Beyond marking an image based on the presence of adult content, the API also returns a hierarchical list of tags with trusted scores. These tags indicate certain categories of adult content, thus allowing granular filtering and management of large volumes of user-generated content (CGU). For example, social and dating sites, photo sharing platforms, blogs and forums, children's apps, e-commerce sites, entertainment services and online advertising services.

• Celebrity recognition - identifies people in users' video and image libraries; can recognize thousands of celebrities in many categories, such as politics, sports, business, entertainment and the media.

• Feeling and demographic analysis - detects emotions such as happy, sad or, surprisingly, demographic information, such as sex, from facial images. It can analyze images and send emotions and

demographic attributes to another service, Amazon Redshift, for regular reporting of trends, such as store locations and similar scenarios.

• Text detection - allows you to recognize and extract textual content from images. The text in the image supports most fonts, including those with a high degree of stylization. Detects text and numbers in various formats, such as those found in banners and posters. In image sharing applications and social applications, it can be used to allow visual search based on an index of images that contain the same keywords. In media and entertainment applications, it can catalog videos based on relevant text, such as ads, news, sports scores, and subtitles. Finally, in public safety applications, it can be used to identify vehicles based on the number of license plates in the images taken by street cameras.

• Enhanced learning image analysis - uses enhanced learning technology to accurately analyze images, find and compare faces in images, and detect objects and scenes in images and videos.

• Scalable image analysis - allows the analysis of millions of images to be able to clean and organize massive amounts of visual data.

Clarifai.com

This image recognition tool allows you to search for images using other images. As redundant as the function of this service may seem, Clarifai has many advantages when a user makes an effort to describe an image to a search engine.

Clarifai offers the ability to recognize images and videos, with a partially free service available to users, based on training artificial intelligence engines to recognize visual content. The API is built around a simple idea: incoming messages (such as image or video) are sent to the service and it returns predictions. The prediction type is based on the user's model. For example, if the data entered by the food model is used, the predictions it returns will contain concepts that the food model knows.

Microsoft Computer Vision

The cloud-based Computer Vision service gives developers access to advanced algorithms for image processing and information return. Computer Vision works with popular image formats such as JPEG and PNG.

To analyze an image, the user can upload an image or specify an image URL. Computer Vision algorithms can analyze the content of an image in different ways, depending on the visual characteristics that interest you. For example, Computer Vision can determine if an image has adult content and can recognize all faces in an image.

Computer Vision can be used to:

• image analysis to detect and provide information about the characteristics of objects in images and the characteristics of images;

• extracting the text from the images;

• image moderation.

Salesforce Einstein Vision

Einstein Vision allows the use of the power of artificial intelligence to recognize and classify images. Pre-determined classifiers can be used or your own custom classifiers can be set to solve specific cases. Images may contain contextual cues about all aspects of the business being promoted, including customer preferences, inventory levels, and the quality of the products being promoted. These cues can be used to enrich sales, service and marketing information to obtain new information about target customers and to take appropriate action.

The possibilities are limitless with applications that include:

• Visual search - expanding the ways in which brand customers can discover products and increase sales;

• Offering customers visual filters - to find the products that best suit their preferences while browsing online;

• Features that allow customers to take photos with promoted products to find out where they can shop online or in-store;

• Brand detection - a sustained activity of monitoring one's own brand on all social channels is necessary to increase the coverage and to preserve the integrity of the brand;

• Functionalities that allow the understanding of the audience's preferences and lifestyle through social media images;

• Features that allow user-generated image monitoring through communities and review panels to improve products and service quality;

• Features that allow the evaluation of banner advertising during broadcast events to lead to a higher ROI (return of investment);

• Product identification - new ways can be extended to identify promoted products to streamline sales processes and provide customers with personalized services;

- Identified product problems before sending a field technician to reduce case resolution time;
- Identified products out of stock or moved to make inventory management more efficient.

4. Conclusions and practical implications for AI Media platform

The new paradigm of cloud processing systems promises online services accessible online based on the model of public utilities services. Like developed companies that want easy access to utilities such as water, electricity, gas and telephony, software developers can use Web Programming Tools (Web API - Application Programming Interface) to increase their productivity. These utilities are becoming more common and meet a wide range of requirements from programmers. The big concerns in the software industry provide such utilities that allow image processing in the Cloud using artificial intelligence and machine learning (ML) algorithms. Convolutional Neural Networks (CNNs) are used to classify and label images, making them extremely easy to use if you use cloud processing APIs. It is known that these adaptive models, especially neural networks, can manage graphic images by interpreting them in vector representations with fixed dimensions.

In order to develop the prototype of the AI Media application, a series of tools available on the market using CNN were identified, which are in different stages of development. However, due to the long waiting time for algorithm training, the use of a Google Cloud Vision API service designed to meet the requirements of the AI Media application was also taken into account.

Google Cloud Vision allows you to identify objects in images, architectural elements and tourist attractions, as well as people or animals. If people are identified, their facial features or emotional states can be differentiated. At the same time, neural networks are involved that can detect popular brands and logos, as well as the products identified in the images. The AutoML Vision feature available in Cloud Vision allows you to drive a neural network with predefined tags, which helps identify new logos or brands that the network has never seen before (e.g. local brand logos).

To make the prototype, the Python programming language and the google.cloud.vision library were used on a test medium in JupyterLab. It allows fast testing in a dynamic way of the API, without the need to compile or troubleshoot the code. The goal is to identify a viable solution to meet the requirements of the project quickly, so later these technologies will be integrated into an application that can be put into production. Cloud vision is an easy interface for the Tensorflow library's Advanced Convolutional Neural Networks package. This means that the AI Media application can have functionality similar to that tested with Cloud Vision in a locally configured system with open source technologies.

Acknowledgement This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0800 / 86PCCDI/2018 - FutureWeb, within PNCDI III

References

1. *Agarwalla, S., & Sarma, K. K. (2016). Machine learning based sample extraction for automatic speech recognition using dialectal Assamese speech. Neural Networks, 78, 97-111.*
2. *Ayodele, T. O. (2010). Types of machine learning algorithms. New advances in machine learning, 3, 19-48.*
3. *Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column deep neural networks for image classification. In 2012 IEEE conference on computer vision and pattern recognition (pp. 3642-3649). IEEE.*
4. *Guimarães, A. J., Araujo, V. J. S., de Campos Souza, P. V., Araujo, V. S., & Rezende, T. S. (2018, November). Using fuzzy neural networks to the prediction of improvement in expert systems for treatment of immunotherapy. In Ibero-American Conference on Artificial Intelligence (pp. 229-240). Springer, Cham.*
5. *Hodnett, M., & Wiley, J. F. (2018). R Deep Learning Essentials: A step-by-step guide to building deep learning models using TensorFlow, Keras, and MXNet. Packt Publishing Ltd.*
6. *Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.*
7. *Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).*
8. *LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.*
9. *Maqueda, A. I., Loquercio, A., Gallego, G., García, N., & Scaramuzza, D. (2018). Event-based vision meets deep learning on steering prediction for self-driving cars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5419-5427).*
10. *Mehta, R. P., Sanghvi, M. A., Shah, D. K., & Singh, A. (2020). Sentiment analysis of tweets using supervised learning algorithms. In First International Conference on Sustainable Technologies for Computational Intelligence (pp. 323-338). Springer, Singapore.*
11. *Price, T. W., Zhi, R., Dong, Y., Lytle, N., & Barnes, T. (2018, June). The impact of data quantity and source on the quality of data-driven hints for programming. In International Conference on Artificial Intelligence in Education (pp. 476-490). Springer, Cham.*
12. *Turing, A. M. (2004). The essential turing. Oxford University Press.*