MEGADAT
big data solutions

Megadat – Applications of neural networks in science

Megadat, d.o.o., is a company with a goal to develop useful solutions in numerous fields of computer science and mathematics. Company projects are tied to the current scientific needs and usually involve applications of different machine learning and data mining techniques to the real-world problems. Company successfully cooperates with different experts from these fields and also puts a lot of effort in educating students and increasing popularity of analytical approaches to data engineering. It is a spinoff of the Institute of Mathematics, Physics, and Mechanics, Ljubljana (IMFM). Here some projects of the group gathered at Megadat mostly in cooperation with IMFM are described.

1. Motivation for the projects of the group

In a modern world, large amounts of useful data are being gathered in all scientific fields including robotics, finance, medicine, advertisement, etc. Big data analysis and machine learning techniques are gaining popularity among researchers and with the help of continuously improving computers, those techniques are used to gain a better insight into various problems and to increase quality of our lives.

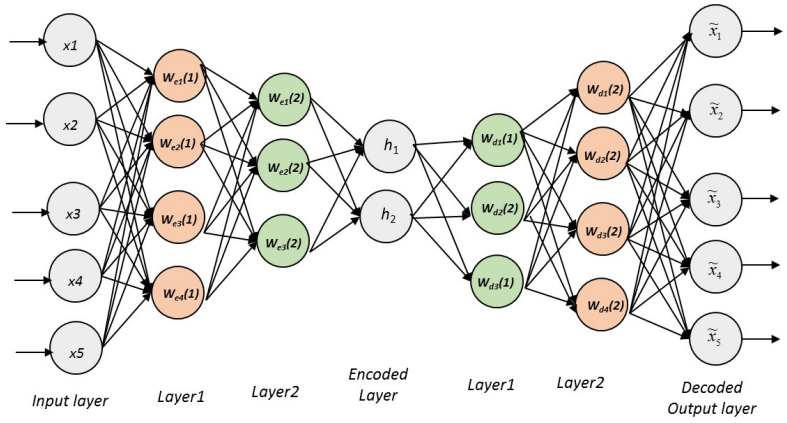

Figure 1: A typical model of a neural network – autoencoder

The main tasks of Megadat's projects are usually designing a machine learning application for solving real world problems and to help extract useful information from the given data. The group has specialized in some specific fields of machine learning including deep learning and has used many kinds of neural networks in group's research. Neural networks (cf. Figure 1) are a special type of algorithms created to mimic the behaviour of the human brain and capable of abstracting feature development. These networks are useful for complex data and are currently widely used by many companies including Google and Amazon. One of their challenges however is their high computational requirements. These complex networks are in demand of a non-trivial computer power such as is provided by high performance computer called Arctur-2 owned by company Arctur.

2. Using neural networks for medical image segmentation

The group has been involved in many projects solving different problems with the help of neural networks. One of the bigger projects was constructing a neural network for 3D medical image segmentation. With the help of publicly available medical images they were trying to construct a convolutional neural network with the aim of locating different irregularities inside human body such as tumors. Convolutional neural networks are designed to work with visual data and have achieved unprecedented performance in the field of image classification, face recognition and similar applications. The group has used data from different patients and their models were designed to help the medical staff with locating the tumor and make better treatment protocols. Problems associated with cancer treatment are among the most important issues faced in the modern medicine and every small advancement can be viewed as potentially lifesaving.

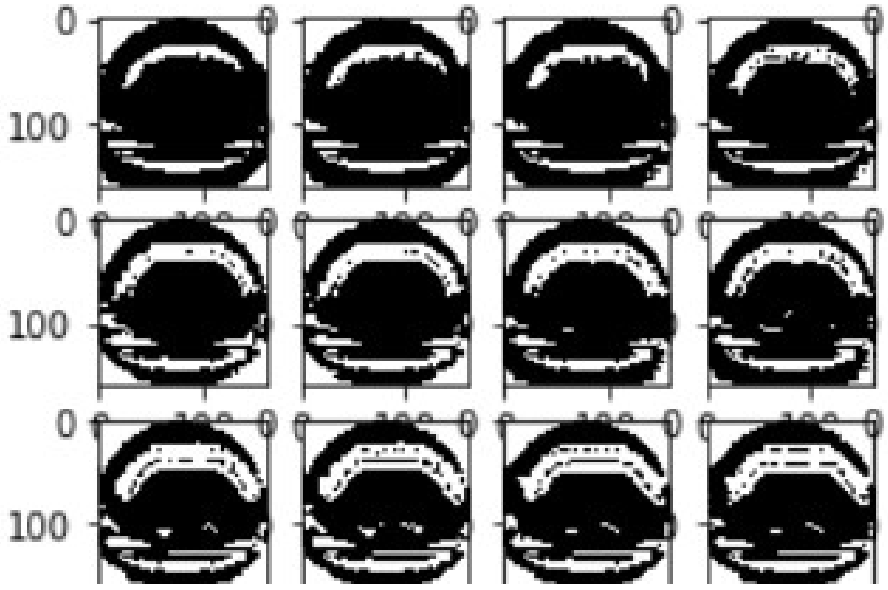

Figure 2: Visual representation of the feature maps used for MRI brain image segmentation

Based on a careful study of the literature on the problem the group chose to ground their approach on the work of Patrick van der Smagt and his research group from Technische Universität München. However, due to the lack of documentation, they had to implement many features by themselves, like a 3D deconvolutional layer, their own activation function layer, different filters etc. With the help of the research group at IMFM, Ljubljana, they managed to get the network ready to use.

They tested their algorithms using publicly available data from the cancer imaging archive containing CT scans of lungs and different MRI images of brains and livers. All these tests were performed on the Arctur-2 computer. Best results were obtained on brain MRI images, where they managed to construct a complex convolutional neural network and locate different irregularities using convolutional filters. Figure 2 presents an example of these filters appropriately visualized as feature maps. Feature maps in convolutional neural networks are a measure for how close a patch of input resembles a feature and they help to identify hidden abstract features in the data, which cannot be manually engineered. The visual representation of these filters can help in understanding the data and features that are useful for the case.

Figure 3: Control panel of the synchrocyclotron at the Orsay proton therapy center, France (David Monniaux)

Automatic image recognition systems have been in the focus of the research in medical science ever since the proton therapy and lately also carbon therapy have come out of the experimental phase and ion beam therapy centers are growing faster and faster around the world. Behind the advertised “targeted and gentle treatment against certain cancer types” there is a serious study of medical images in highly equipped diagnostic centers. The methods behind the scene are often various machine learning approaches with deep learning becoming one of the most promising ones. As opposed to the standard radiation approach the newly developed beams can hit the tumor as hard as possible while keeping the damage to neighboring healthy cells to a minimum. This brings into the spotlight the question of accurately computing the right spot – location of the tumor – in the diagnostic phase. Some therapy centers such as Heidelberg Ion Beam Therapy Center can afford their own big computers in order to make their approach really highly effective yet safe as promised. Smaller ones may rely on the supporting computer power such as the one offered by Arctur.

3. Other projects – Zemanta

Aside from medical research, the group has successfully used neural networks and similar techniques with many different types of data. One of their engagements was with the company called Zemanta, whose main product is an enterprise platform for content marketing. The main goal of the project was to develop a machine learning algorithm to predict a probability of a user to click on the proposed link representing a certain web advertisement. Zemanta allowed them to train the group's favorite methods on their original data. Different approaches including neural networks were tested and compared with the aim of choosing the best possible method for this specific problem. They managed to train a suitable model and to compare the results with the models that Zemanta is using. Neural networks along with some other methods showed promising results, but for further development it is necessary to incorporate the solutions into their existing structure.

They were also involved in some projects in robotics, where they managed to use different deep learning techniques for robot training and simulations. All research was conducted in the Institute Jožef Stefan in Ljubljana, where they were given access to their robotic arm and simulators. The main objective was to help the robotic arm to learn simple movements faster using neural networks. They managed to develop and train a special neural network called Autoencoder, which was used in combination with other machine learning approaches. The inclusion of neural networks showed the desired results and improved learning rates of robotic arm on a number of tasks. The results showed the possible benefits of using neural networks in robotics and hold a significant scientific value. Their work was continued by the IJS specialists.

4. Conclusion

Megadat will continue to work on these tasks, especially on automatic segmentation of medical images and some other problems arising in ion beam therapy. They also intend to widen the diversity of their products and start cooperation with some other firms including Cosylab.