

gcForest

Marcel Aguilar Garcia, ID: 20235620
m.aguilargarcia1@nuigalway.ie

25st March 2021

1 Introduction

Deep Learning has recently gained a lot of popularity and the usage of Deep Neural Networks has seen an increase in many applications such as image and speech recognition. However, while DNNs outperform other machine learning methods in many applications, their complexity has brought some challenges to scientists. In general, DNNs have a large amount of hyperparameters and they are known to be very sensitive to them. Additionally, DNNs complexity makes their interpretability complicated and they are usually known to be black-box. Finally, DNNs do not usually work good with small data sets and a large amount of data is required to achieve good performance.

Until now Deep Learning has always been seen as part of Neural Networks but recent studies have found alternatives using non-NN models in order to build deep models. Zhi-Hua Zhou and Ji Feng have introduced Deep Forest which uses layer-by-layer processing architecture from DNNs while using Random Forests. Deep Forests have less parameters and they are more robust to hyper-parameters settings. Additionally, Deep Forests configuration can be determined in a data-dependant way which makes them more suitable for small datasets.

As seen in *paper*, each Deep Forests layer contains several random forests, completely-random tree forests, etc. Given an example, each forest will return the class distribution for that example. The average class distribution across all forests will be computed within that layer and passed to the next layer. Using this idea the authors extend gcForest to use multi-grained scanning. In the final results, gcForest was able to achieve very good results and compete with Deep Neural Networks.

While Deep Forest is able to provide very good results with the original structure, the author believes that by extending the vectors that forests returns the accuracy will be able to improve. In order to understand the class distribution that a forest returns we first have to understand what this class distribution is in a tree.

By default each forest returns the class distribution of the instance. For a classification problem with N classes, let $\mathbf{C} = (c_1, c_2, \dots, c_N)$ be the set of the different classes. The current architecture for Deep Forests makes each forest return $P_{Forest} = (p_1, p_2, \dots, p_N)$ where p_i is the probability of the class c_i being predicted by each forest. Let's take a closer look to how this value is achieved.

2 Research Method

This section may include comments of what search terms you used during your library and web search e.g. The search terms used were "Artificial Intelligence" AND "Accident" OR "Emergency Departments". You should also say where you searched e.g. "The author searched for relevant references from the NUI Galway Library, Scopus, Google Scholar and Google in general." You might also make reference to the dates of the articles you selected e.g. "Articles older than 2010 were excluded from this research except where they were of historical importance".

3 Other sections...

These sections cover main results of your literature search and are densely cited.

3.1 Key Elements to Include in Your Report

You must include the following elements in your report:

- Introduction section
- Research Method section
- Other sections e.g. consider your research questions
- Conclusions section
- List(s)
- References (at least three)
- Subsection(s)
- Original Figure or infographic

4 Conclusions

Your conclusion is the only place where you can express your own opinion on your hypothesis based on what you found in the literature e.g. did you find good evidence to support or contradict your hypothesis?

A Sample Hypotheses

Choose your own research topic or be inspired by one of the hypotheses below.

Spam will cripple our email systems in the future ; World Wide Web is changing our definition of copyright ; Mobile computing will make significant changes to society ; Open source software makes an impact on society ; Identity theft will become a significant problem in the future ; Social media is changing the work of political activists ; Google users have no legitimate expectation of privacy ; Apple's philosophy on closed systems is correct ; Driverless vehicle technology creates great value for society ; IT has helped change the role of libraries and the activity of research ; The World Wide Web has added to the problem of academic plagiarism ; Information technology reduces the risk of medical errors in hospitals ; Telemedicine a realistic way of providing medical care ; IT technologies and the Web has adversely affected sleeping patterns ; Stricter copyright laws are required to protect the film and music industry ; The Web is now more effective than books in primary schools ; Women have equal opportunity in the IT industry ; Online and blended learning is more effective than traditional classroom learning ; Children can be taught to teach themselves using the web ; The Digital Divide adversely affects economies ; Food distribution and traceability are greatly enhanced with computer technology ; IT technology adds value to the arts (such as drama, dance, literature, music, visual art) ; Advances in graphics, networks and artificial intelligence have changed video games ; Reliability of health and medical information available on the Web is unprotected ; Intellectual property (IP) is being changed by the Web ; Facebook's privacy settings are bad for users ; RFID tags have revolutionised sales of clothing ; Wearable computing presents more problems than solutions ; Augmented Reality presents more risks than opportunities ; Tablet devices will replace books in all schools ; Virtual reality is the next wave of information technology ; Describe the current state of gesture based computing. ; Describe the current state of visual data analysis technology. ; Digital Immigrants and Digital Natives are becoming more widespread ; Web 3.0 is closer than we know ; The social semantic web is here ; MOOCs are the future of higher education ; User interfaces are about to move to the next level ; Hacking groups provide value for society ; Data mining presents significant challenges for society ; Hawkeye has been great value for GAA sports ; Electronic Medical Records hinder communications between physicians and patients ; The Internet provides support for people living under oppressive forms of government ; Twitter and other forms of social media adversely influences public opinion ; Ireland has not embraced the opportunities of eGovernment ; Off-shoring and outsourcing provide greater opportunities for poorer economies ; Computer technology has adversely impacted waste use and disposal ; Facebook determine peoples mood ; Bullying is now a larger problem because of social media ; People with disabilities have not benefitted from computer technology ; Social media is now widely used for

marketing purposes ; Smart homes are here ; Computer software in our lives is now 'safety critical'; The 'dark web' is a concern to governments world-wide ; Computers will soon perform better than humans ; Computer addiction is real and should be a recognised disorder ; 3D printing has positive outcomes for users ; The Agile method is worse than the Waterfall method for software development ; Web page speed: The technology and psychology; its effects on website and user behaviour. ; Technology is training us to become more distracted ; Society will soon no longer use cash or cheques ; Image recognition systems will soon make small changes to images that humans cannot notice ; Technology improves the independence of the elderly ; Social media will soon be replaced by something more sinister ; Consciousness will soon be possible in computers ; The internet of things (IOT) will make society worse ;