

[Inspiratron.org](https://inspiratron.org) - [Natural language processing, machine learning and cybersecurity](#)

Ideas for the future

by Nikola Milošević? - Tuesday, December 11, 2018

<https://inspiratron.org/blog/2018/12/11/ideas-for-the-future/>

I would like to state a couple of ideas that I have been thinking in the past number of days regarding what I do, which generally is natural language processing and machine learning. They may be something I am already working on and some ideas for the future and future directions. Only time will tell which I will manage to tackle.

Named Entity recognition

Interesting topic, however, it seems it is moving more towards industry domain, rather than academia. However, there are still a lot of people working on this, especially in specialized domains, such as biomedicine. Lately, we are having at the University one big project related to anonymization of clinical health records and lab reports. As a first step in the process is named entity recognition. Also, a lot of shared tasks such as TAC, N2C2, on which I have participated with some people from the school are focused on named entity recognition. However, some companies, including Amazon published tools, such as Amazon Comprehend that are tackling the issue of named entity recognition. So has this become an industrial task now, rather than academic? Maybe. However, there is still a blurry line, and both industry and academia are working on NER problems.

Deception

Deception is when someone tries to manipulate the situation in their favor using partial truths, lies, leaving out information or under- and overrepresenting some facts. It can be a very severe issue in many aspects. Using deception people can gain an advantage over other people, they can mislead people into financial losses or induce other damages for them. It can be a severe issue from the security point of view, as certain deceptions and frauds are considered a crime. Militaries for ages used deception to gain an advantage over the enemy. Therefore detecting enemy deception can make advantage in a military sense. Financial market and shareholders of companies may be victims of deceptive activities of some of the market players. Human communication relies on truthfulness, however, that is not always the case and deception detection may reduce significantly some of the risks that we are facing at the moment.

It has been reported already that people don't like to lie or use deception, and therefore in some minuscule way, they would change the way they use language. For example, they would use more often passive voice, they would change some of the pronouns they use in order to distance themselves from what they are talking about (talking about you, they, them, instead of we, me, us), etc. Machines and machine learning is quite good at detecting these things, either as an anomaly in someone's use of language or as a classification task. I have thought of the following research questions as something worth researching:

What language feature, such as frequency of determiners, tenses, passive/active voice, use of pronouns can determine that someone is trying some deception technique?

What is the role of context and background knowledge in deception and deception detection?

Can deception be modelled as anomaly detection task? If people use language differently when using deception, and if most of the time people are truthful (are they?), then deception would cause an anomaly in someone's use of language.

Transfer learning in natural language processing

There has not been so much active research on the use of transfer learning in NLP. I guess it may be due to transfer learning being a relatively new area, but it is getting hot. In my opinion, NLP can greatly benefit from transfer learning, as from domain to domain there are just slight differences, while the majority of concepts may remain the same. I am even starting to believe that in transfer learning may lie a key for more general artificial intelligence. However, we are probably far from achieving it yet, but if you combine deep neural networks, reinforcement learning and manage to get some reward/punishment system in place so you have continuous learning through reinforcement learning, that can be a step closer. Like initial learning is done using deep learning and standard learning procedures, then the transfer learning can apply it to more general area and reward and punishment mechanism would improve it over time.

However, let me go back to the are of NLP and transfer learning first, before expanding the idea. I think some of the research questions can be:

Ideas for the future - 12-11-2018

by Nikola Milošević - <https://inspiratron.org>

Which NLP and in what way can benefit from transfer learning?

Can similar concepts (e.g. city vs country, doctor vs patient) benefit from transfer learning and use each other to learn some features of each of them, improving performance?

What are different techniques to be used to transfer knowledge?

Memory in learning and can we use it to expand on transfer learning?

These are some of the ideas so far. I will probably try to write some proposals along these lines, for some grants and maybe even some fellowship application. I still don't know, they are cooking in my head. I believe they are quite good areas of research and may be successful. This is the first attempt of writing down what has been in my head these days. I am keeping it only for my patreon supporters so that it is a closed circle of people before something actually happens. However, in future, I may be writing more content that may be limited to patreons only, but as well publish very public posts.

All rights reserved and copyrighted by inspiratron.org and Nikola Milosevic