Natural Language: The Road to Infinity

Jane Nemcova - August 17, 2021

Abstract

- While the earliest applications of AI involved "language" and showed great promise, the latest AI innovations have revealed what many of us in the language industry have known all along: "language" is complicated.

- Natural Language has intrigued and baffled the AI world, revealing that the expertise required is infinite. Whether as consumers, creators or citizens— language highlights the importance of human knowledge in developing AI, where the future hinges on our preservation of human language, culture, and individuality.

- AI has now squarely taken root into every aspect of our lives. But, will machines eventually (inevitably?) replace us as linguists? The many evolving approaches in AI guarantee that this last mile towards optimizing machines will remain just that... the ever-expanding "last mile".

- As such, we should turn our focus towards the needs of the market as well as our obligations to promote the human-side of language. Not only is the expertise critical for business success, it is desperately needed for the future of our unique cultures and countries.

An AI Future

- Expertise in language translation

- In AI, Language is the most complex area to master

- We have skills that make us different types of thinkers and experts

Business World Understands This

- *"What looks like a great job graduating from college today might not be a great job five years from now."*

- *"I personally think there's going to be a greater demand in 10 years for liberal arts majors than there were for programming majors and maybe even engineering... When the data is all being spit out for you for you, options are being spit out for you, you need a different perspective in order to have a different view of the data." Mark Cuban*

AI
What do we mean by "Intelligence"?

- Intelligence has always been the domain of humans

- Intelligence is intrinsically human. An examination of intelligence, in its essence requires a study of human nature and understanding its unique qualities and aspects with the natural world

Deep Learning

- Deep learning - while extremely effective under certain circumstances - has its limitations, particularly in language

- Language is a reflection of cognitive reasoning (human intellect)

- While we utilize logic and other rules and patters in language, there are always exceptions and innovations which cannot be learned by machines automatically.

AI in the Marketplace

- The concepts of AI have been around for decades

- Only in the last 7 years has AI received so much attention in the broader marketplace

- AI is influencing every aspect of our lives, as consumers, citizens and business people

AI v. Technology

- "AI" has almost come to be synonymous with "technology", but these are different concepts

- While we might not understand well AI as consumers, we feel it's presence and see its business evolutions in companies

- While AI is rolled out into the marketplace, it gives rise to societal and ethical considerations

Natural Language

- Our ability to understand and make sense out of human language becomes a very important value for intelligent systems

- If done correctly, the components of NLP become essential for whether an AI system can ultimately perform its intended goals

Why is NLP Important?

- NLP is considered to be the most difficult component because machines are not able to "understand" as humans do

- How language is created, generated and understood in humans is something that philosophers and linguists have spent thousands of years studying

- However, it is still not entirely clear how the mind "thinks" and thus how language is generated

- Language both reflects concepts but also creates opportunity for ambiguity in "meaning"

Math v Language

- Language is not univocal

- While mathematic representations provide great utility in building intelligent systems around language, language is unique in that meaning and form is not represented in units (quantities) that reflect fully an idea that a human generates

- We see this challenge every day in the translation and localization field

- Language and meaning are the very challenging nuances we try to address as translators

Language & Meaning

- Language and meaning are prone to ambiguity

- We see this in normal human-to-human exchanges, where even when the language barrier is not an issue, humans still have a "meaning" barrier or do not share similar "concepts"

- These factors provide opportunities for ambiguity on a variety of levels

Further Learning

- Examining the origins, structure and correlation of language as it relates to the more complex aspects of reasoning, cognition and the soul

- At a language level (vs meaning and concepts), we need to keep in mind that intelligent systems need structure and context in order to start processing natural language

- Therefore, to be successful in NLP, we must try to make language less ambiguous for that machine

Further Learning

- Deeper studies in lexical and syntactic analysis, semantics, phonology, morphemes, morphology, pragmatics and other linguistic areas

- As translators and localizers you are well aware of these nuances in language, and are able to understand the principles of such thinking and analysis in order to see how language needs to be structured and organized for an intelligent system to understand.

NLP Applications

- With the rise in AI products, NLP is more crucial for its products' success. As AI has evolved, so have NLP approaches

- While the uses and applications of AI are increasing, so are the tools used to perform NLP, resulting in better and better performance - which is evident in all AI applications we readily encounter on our phones, home devices, chatbots and cars

- Though there is debate about the effectiveness of approaches in language, it would be important to note that NLP methods have in practice been making use of deep learning more and more frequently to solve language challenges

Importance of Data

- "The analogy to deep learning is that the rocket engine is the deep learning models and the fuel is the huge amounts of data we can feed to these algorithms." Andrew Ng

- Using a deep learning approach essentially requires good algorithms. At the same time, these algorithms are dependent on the data that is fed to them

Data Outsourcing

- As the most important AI companies started developing these DL models, they quickly realized that they needed to start leveraging the data that they already held within their companies

- Most companies - in realizing how they would approach this - discovered that they needed organized and different data

- This generated a large demand for outsourced data support services, particularly where human scale and coverage was something companies lacked infrastructure for

- This has lead to large scale data outsourcing partners

Data Bias

- In its essence, all data is intended to bias an intelligent system towards learning

- It is helpful to think of algorithmic bias and data bias as two separate topics

- Both are intended to improve the decision-making elements in AI, though imperfectly

Good and Bad Data

- AI and machine learning use quantitative elements to represent ideas in order to make human-machine interfaces functional.

- By reducing complex ideas and concepts, we tend to limit idea representation

- While this means that we can utilize data for training AI, it is important to note that we are not always able to represent conceptual complexity adequately and appropriately in machine learning approaches

- Important to know not only what data is collected, but how it is collected and it's parameters and scale

Subsequent Questions

- What is the impact of bias in Machine Learning

- What to do about bias in AI

- Why is Ethics so important all of a sudden

  - We see issues arising as AI moves from the labs into the real world, which gives rise to further questions we should always ask when tackling AI ethical questions:

  - What do we use the AI for?

  - Who has access to it? Who controls?

  - What are the potential outcomes we expect and those we don't expect?

New Ethical Considerations

- AI becomes quickly more tricky in practice

- Organizations will need to strive for transparency, inclusivity, explainability and alignment

- While humans are very intelligent, we must couple this knowledge with humility, as we do not know everything, in all ways, and in all appropriate contexts

- Language becomes a focal point of ethical considerations across the board

New Opportunities

- What potential career directions can I take in AI?

- What are the directions in the market with the most demand for people with language skills?

- What might be relevant future issues to address?

There will be a focus on Language

- Traditionally, we tend to think that engineering is the most important career pursuit in AI - however, the more that technical capabilities and tools improve, there will be increasing focus on the roles that other areas of expertise will hold in making the AI field adapted to statutory, ethical and employment considerations.

- The most challenging aspect in AI involves linguistics - in the domains of speech, text and image, the most complex area is language.

- Therefore, language experts will be in higher demand and bring a skillset and perspective that is unique.

Variety of Roles

- Speech Analysts
- Raters
- Graders
- Annotators
- Computational linguists
- Translators
- Language engineers
- Data Scientists
- Data Analysts
- Data engineers
- MT engineers and experts
- Researchers
- Speech and Data Program Management

Roles Across Functions

- With machine learning jobs rising in areas such as natural language processing and deep learning, the number of possibilities are extensive regardless of the specialty you choose to pursue

- Furthermore, machine learning is poised to impact every industry, and your particular field will eventually need such experts, where the furtherance of your understanding of the principles of language and AI will help you to seize opportunities in the automotive, banking, customer support, sales, entertainment, education, and every industry where you may have interest to work

New Era in Language

- Language is such an important part of our humanity

- Our ability to speak, create and understand language is our closest participation in spirituality, defining our humanity

- Those with such knowledge need to adapt to the market and are essential in bringing this knowledge to AI, in an infinite number of ways