Characteristics of AI-infused systems

Main characteristics of AI-infused systems

- Learning
- Improving
- BlackBox
- Fueled by large data sets

General machine learning (Classical Learning overview)

Learning and improvement process in an AI learning process mainly involved in general machine learning approach that is a system make an action base on learned background that based on large dataset.

Machine learning is a subfield of study in artificial intelligence that enable machines to learn autonomously without explicitly programmed.

In 1959, Arthur Samuel defined ML as [1],

"*Field of study that gives computers the capability to learn without being explicitly programmed.*"

ML, provides machines the ability to learn based on experiences, observations and analysing patterns in the given data. Its goal is to understand and follow the instructions based on the algorithms to perform the task automatically without any human interventions. The basic premise in ML is to build algorithms that can receive input data using statistical analysis to output more accurate, in predicting outcomes.

A typical example of ML Black-Box

A well-defined ML as follows [2] [2]:

A machines is said to be learning from **Experience** (E) with respect to **Task** (T) and **Performance** metrics P, if it performance at tasks in T, as measured by P and improves with experience E.

Then, $E * T = P$

Based on **Experience** E is placed against **Task** T for improvements, and measured by **Performance** metrics P.

Experience: Input data	Task	= Performance
Emails	Spam detection	Block unwanted emails
Images	Categorize and recognize images	Organize images
Transactional data	Segment user data	User classifying

Example: Playing Go

T = Playing Go

E = Gaining experience by playing practice games of GO

P = the probability of winning a game against an arbitrary opponent

Data collection

However as described in (*Guidelines for Human-AI Interaction*)[1], that an AI infused system may react differently depending on dataset continuously evolving and changing overtime, and these inconsistencies and unpredictably might confuse users, erode their confidence and lead to abandonment of an AI technology.

In typical ML process, to prepare data for machine learning initiatives one can accelerate machine learning projects to deliver an immersive business consumer experience that accelerates and automates the data-to-insight pipeline by following varies critical steps, such as standard data collection, profiling and splitting data into training and evaluation sets. This is done through streamline data from different source system that could an application portal via database or web-services.

A typical example that might reflect characteristics that given above in an AI-Infused system is voice recognition by virtual assistant like SIRI or Amazon Alexa. For example, if we set our IPhone any handheld devices in a language that is other than English or set to English but the user is non English speaker.

Then it becomes obvious that virtual assistant struggles voice recognition imbalance in the AI system. This problem is well-known in interaction with people with disabilities, it has been proven that Google speech recognition system does not work well for people who are deaf and hard of hearing.

This is a typical scenario, where system pre-learned on data that cannot handle real live issues. Where dataset that system has been trained on tackle a standard scenario but did not included in the data model set.

Take for example, SIRI on Norwegian language mode and only trained on dataset based on people from eastern region of Norway where people have certain type of dialect and the system is being used in southern or western region of Norway then voice recognition would face challenges to understand the end-user. This type of problem real live consequences in users who has speech impairments where voice recognition system is a part the smart home system [5].

Human-AI interaction design

Design guideline G5 and G7, these guidelines fall into the characteristics that been used in previous task.

G5 AI design guidelines is the guideline during interaction ensures the experience is delivered in a way that users would expect, given their social and culture context. This guideline falls directly into the problem that has been described above that if voice recognition system has been trained on certain dataset and struggles to understand user who uses different type of dialect or has speech impairments then example guideline propose application guidelines such as use of semi-formal voice during interaction with the user and use word validation at the end of every dialog with further question assistants. This will make voice recognition system more durable since semi-formal voice brings AI to a human level day-to-day interaction that could lead to use of simple set word that may implicate set of instruction and validate those instructions end the of every interaction.

Disadvantage with this guideline may lead to problem with people who has speech impairment where resulting in atypical and relatively unintelligible speech in most cases. Some characterizations of these speech impairment by a slurred, nasal-sounding or breathy speech, an excessively loud or quiet speech, problems speaking in a regular rhythm, with frequent hesitations, and monotone speech. An extreme case of these speech impairment is Dysarthric (see [5] under Dysarthric Speech for VA).

G7 AI design guideline is the guideline when wrong that support efficient invocation, that is make it easy for the user to invoke or request the AI system services such that user can a particular command to initiate. In smart home system where voice recognition AI can initiate a given task just by a command or a set of commands to complete a particular task or tasks.

In this scenario where a user struggle with speech impairment could easy get to initiate task only using task command rather than use a set of dialog with the AI system, where AI could struggle to understand the end-user because deviation in voice input by the user. This type of command implementation makes the AI to sharpen to focus on a simple sentence or word rather than to process through a large string of words or character where AI has match diagnostic based on previously trained data model, which may contain inconsistency.

Simplicity in this process, a manufacture could still pre-train an AI beforehand before put in a production environment where making errors during runtime would be minimum rather than a trained AI with fully speech recognition on incomplete dataset and that requires the AI to learn over longer period of time and eventually alteration in UI input which might end up with discouraging the end-user to use the AI system altogether.

As recommended in (*Guidelines for Human-AI Interaction*)[1], that building in a safeguards like verification steps or controlling levels of autonomy to help prevent unwanted adaptations or action from intelligent systems. Which goes back to inconsistencies and unpredictability in using an AI-infused system where user had to struggle to adapt as well how AI evolving over time.

Chatbot / conversational user interfaces

In luger and Sellen (2016) key challenges revolves around the human and machine reliance, that is interaction mostly becomes who the users are and their prior background. In luger and Sellen user profile was most dominating aspect whether a user would be effective in using a AI conversional agent (CA). In some instance where they show a cultural implication that motivates the users trust in AI to completing a task rather than make trust-decision solely based on the AI capabilities. So challenges are very dynamic some part of it based on user and machine interaction evolves over time, in some cases in good direction and in some otherwise. Challenges that have been presented in Luger and Sellen are some of the same problems we have discussed in the previous sections.

Some of the key challenges that have been mentioned are:

Uncertainties in system capabilities, such as how competent is the CA system, which lead user to believe that some interaction to complete certain tasks limited and could have been done more efficiently if they knew intelligence limit of the CA system. The other side of this problem was users did not know if the system has learned from their interactions over time, that might imply if the CA has altered its capabilities.

Users wanted some form of validation from the CA system such as learned or know, or this could initiate as a learning process for the CA as user by themselves could initiate learn command to the AI to learn from the interaction or build up its knowledge base for future interaction where AI could automatically make suggestions to the users. Basically, what users were recommending are system feedback loop where user can personalize the CA system based on their preferences.

Assessing system intelligence, In terms of perceptions of CA intelligence, This resulted in the majority of users being unsure as to the interaction dynamic; is the computer learning to adapt to the user or visa versa. However, for those with lower levels of technical knowledge, a combination of (a) the system's failure to learn/adapt to either their accents or the ways in which their questions were posed, and (b) its tendency to resort to web search, led them to frame the CA as simply a voice-based search engine extension. These lead many users to believe that a CA system is just a voice recognition system that convert speeches into word string and uses web search engine to only look up the information that has been queried but not making an intelligent decision based on the commands.

Users expected the CA to be able to infer, from all previous interactions, the context of the current task. In particular, once an interaction/task was complete, the majority of users expected the CA to remember the context of the preceding interaction. Equally, more positive ‘conversational’ experiences were reported when the CA was perceived to have understood the context of use, for example knowing that reading a message would likely lead to the user wanting to reply.

AI design Guidelines G1, make clear what the system can do. Help the user understand what the AI system is capable of doing. This guideline preciously tackles the problem of systems capabilities that is present all the metrics it tracks and present it to the users. Metrics that involves data it has been tracking over time by interaction and previous choices that have been made by the users categorised by application or search commands and results.

AI design Guidelines G2, make clear how well the system can do what it can do. Help the user understand how often the AI system may make mistakes. As it mentioned in assessing system intelligence where users expect system to come up with suggestion based on user preferences that AI can infer suggestions based on previous choices.

Bibliografi

[1] A. L. Samuel, «Some Studies in Machine Learning Using the Game of Checkers. II -Recent progress».

[2] A. B. Amal Kumar Choubey, «GeeksforGeeks, a computer science portal for geeks,» GeeksforGeeks, 2019. [Internett]. Available: https://www.geeksforgeeks.org/ml-machine-learning/. [Funnet 2019].

[3] «ZLD,» ZLD, 4 2017. [Internett]. Available: https://www.zldoty.com/what-is-machine-learning/. [Funnet 6 2019].

[4] D. M. V. A. F. B. N. P. C. J. S. S. I. P. N. B. K. I. J. T. R. K.-G. a. E. H. Saleema Amershi, «Guidelines for Human-AI Interaction,» Microsoft and Paul G. Allen School of Computer Science & Engineering University of Washington, Redmond, WA, USA and Seattle, WA, USA, 2019.

[5] F. C. a. L. D. R. Fabio Ballati, « "Hey Siri, do you understand me?": Virtual Assistants and Dysarthria,» 2018.