Fine Tuning Modeling Through Open AI

Chapter · June 2021

CITATIONS
0

READS
401

2 authors:

Varsha Desai
V.P.Institute of Management Studies & Research, Sangli
18 PUBLICATIONS **13** CITATIONS
SEE PROFILE

Kavita Oza
Shivaji University, Kolhapur
99 PUBLICATIONS **175** CITATIONS
SEE PROFILE

ISBN – 978-81-948061-2-7

Progression in Science, Technology and Smart Computing

Fine Tuning Modeling Through Open AI

Varsha P. Desai

Research Scholar, Computer Science Dept. Shivaji University, Kolhapur

Dr. Kavita S. Oza

Assistant Professor, Computer Science Dept. Shivaji University, Kolhapur

ABSTRACT - Open AI is an artificial intelligence research laboratory working on cutting-edge deep learning techniques that allows computer to works like a human being and help to solve complex problems. Natural Language Processing is one of the applications of deep learning through Open AI that motivates to train billions of parameters with complex and large corpus dataset with better performance. GPT-3 is auto regressive language has capability to convert text to image, face recognition, handwriting recognition, translation, sentence analysis, intelligent recommender. Responsive AI system uses behavior analytics for faster corporate decisions. This paper elaborates the fine-tuning applications of deep learning model through GPT-3 under open AI system.

Keywords: GPT-3, NLP, Deep Learning, ROUGE, Responsive AI

INTRODUCTION

Today deep leaning and neural network are the power of industries. Open AI is an advanced more human centric intelligent technique especially in field of reinforcement learning. It is artificial intelligence laboratory where research scientists explore their knowledge and skill for innovations in machine learning techniques. Artificial intelligence is the technique to allow the computer to behave like human being. Deep learning algorithms works like human brain that analyze complex data with huge logical combinations.AI has competency to enabled deep learning model to solve complex mathematical problems better than machine learning. Deep learning models integrated with automatic optimization of feature extraction process rather than machine leaning model.[1]Multilayered Neural network implementation using deep learning algorithms promotes more accuracy in prediction of complex problem results. Open AI technology motivates for creating image from text, connecting image to text, text analysis, language recognition, microscope etc.

Artificial intelligence is the simulation of human intelligence into machine for thinking and working like human that has capability to work form experience. Machine Learning is a part of AI that has an ability to learn from data, identify pattern and take decision with minimum human intervention. Deep learning is the subset of machine learning in which multilayer neural network learn from large corpus dataset for making intelligent decisions.

ISBN – 978-81-948061-2-7

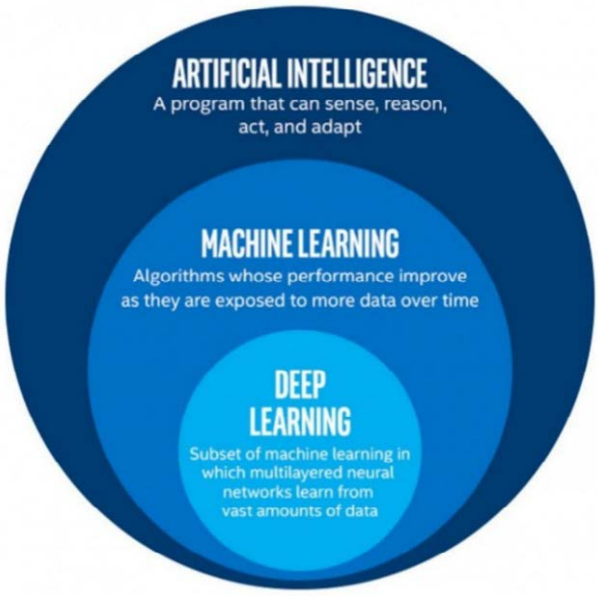

Fig.1 Cousins of AI [1]

GPT-3

Generative Pre-trained Transformer 3 (GPT-3) is an autoregressive language model to produce human like text using deep-- learning algorithms. GPT-n series created by Open AI in the research laboratory. It is one of the effective technologies for Natural Language Processing(NLP).Autoregressive model has a capability to predict the outcome by random processing. Today it is a challenge to work on many NLP data and training the large corpus dataset by fine tuning on particular task to produce results. GPT-3 is a deep learning model with 175billion parameters, more than 10 thousands non sparse language model with fast and efficient test performance.[2] It provide strong performance on many NLP datasets like translation, comments, forums, cloze activity tasks, domain adaption and reasoning data analysis.[3]

Integration of Figma plugin and GPT-3 used to create interactive templates. It automatically generates code as well as provide comprehension of code written by programmer in python. It has capability to generate (JavaScriptXML) JSX layout from plain English text. Using GPT-3 we can generate regular expressions from different use cases written in plain English sentences. By combining capabilities of Figma and GPT-3 use to generate clone website from existing URL. It is intelligent techniques which tell us what things can be done with the inputted object. GPT-3 has a power to generate Automatic chart and plots from plain English. It can be used in interactive quiz designing for developing personalized eLearning applications. Learn from anyone is the GPT-3 tools where learner can select expert is particular area and to get knowledge from them by just typing query in plain English. GPT-3 is can write ML model for specific dataset and has capability to generate code for ML model only through dataset explanation and required output. It work as intelligent recommender,

Interactive Voice Response (IVR) provider, automatic resume creator without training and massive data uploads with greater accuracy[4]

APPLICATIONS - OPEN AI

Jukebox model is developed using deep leaning technique to generate raw audio in different styles and artists. Hierarchical Vector Quantised-Variational Auto Encoder (VQ-VAE) algorithms compress the music into tokens. This model generate the song piece of multiple minutes long which can recognizable singing in natural language voices.[5]

Reinforcement learning(RL) used to handle complex task by human judgment to provide positive and negative comments to the task, summarization of task. Human labels are used to train the models of reward and optimize the model. Deep learning algorithms are implemented to learn from human interaction. RL fine tuning of language model is developed for NLP tasks like high sentiment analysis, CNN/Mail summarization, TensorFlow (TL;DR) dataset. Both supervised learning and reward based learning algorithms provide better results for the NLP task. Communication between human and ML Model gives scalable reward learning methods such as amplification, debate and recursive reward modeling.[6]

GPT-3 is an autoregressive language has an ability to work on large corpus dataset with billions of parameters with better test performance. It achieves strong performance on large number of NLP datasets for translation, cloze task, question-answering that required on-fly reasoning or domain adaption like 3-digit arithmetic, finding novel words in the sentence, unscrambling words. In Fine-Tuning approach supervised dataset is trained by updating the weight of pre-trained model until it results better performance. In few shot approach few demonstration of the task with inference time condition without updating weights. In one-shot approach model assign a single task at a time without updating gradient.[7]

Responsive AI system development ensures that system provides benefits to the society without any harm or negative impact. It involved testing the safety and security of system during development by identifying structural risk associated with AI system. AI security is very important to protect from being attacked, misused by bad actors and co-opted. In real time AI application sometime it is difficult to internalize harm from AI system like social harm due to increased use of AI system may reduce trust in online sources. AI targeted regulations includes government regulations, international standards and clarity on applying existing law to AI system. Responsive AI system development is an integration of action problem that expect benefit from being equilibrium. AI companies are facing many actions problems like more confidence that other will cooperate, assign higher value to the mutual cooperation, low expected cost, assigning low expected value to non-reciprocating cooperation, assigning lower expected value to the mutual defection.[8]

Human Feedback Summarization: Trained the model to prefer human feedback summary by using reinforcement learning. Extensive analysis use to understand human feedback dataset and fine-tuned model. Recall Oriented Understudy for Gusting

Evaluation (ROUGE) is the technique used for automatic summarization of text and machine translation of human produced text. Our reward model is developed that result in better summaries from human dataset. Reddit-trained human feedback models also generate high-quality summaries of news articles on the CNN/Daily Mail (CNN/DM) dataset without any news-specific fine-tuning, almost matching the quality of the dataset's reference summaries. Optimizing our reward model directly results in better summaries than optimizing ROUGE according to humans.[9]

Due to industrial revolution Artificial intelligence demand is increasing day by day. These include financial services, defense, consumer retail, advertising, entertainment. According to market research report it is forecasted that global revenue of artificial intelligent product and services will grow to $36.8 billion by 2025.[10]

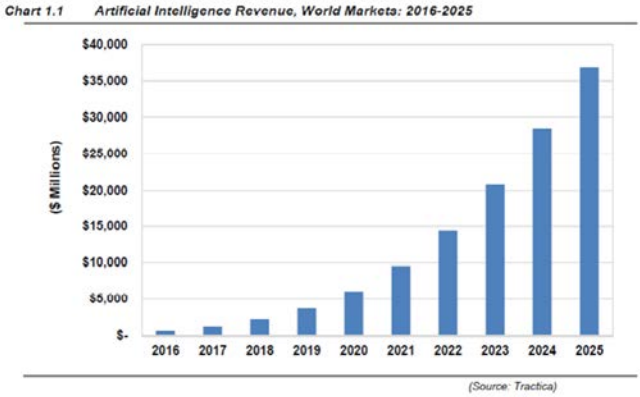

Fig.2 AI Revenue Growth World Market 2016-25[10]

According to market survey repot of Tractica Artificial Intelligence revenue for top 10 use cases of AI by 2025 generated form applications like contract analysis, Object detection and classification, Automated geophysical featuredetection, text query of images, content distribution on social media, predictive maintenance, processing patient data ,image recognition, classification, tagging, algorithm treading strategy performance improvement.[10]

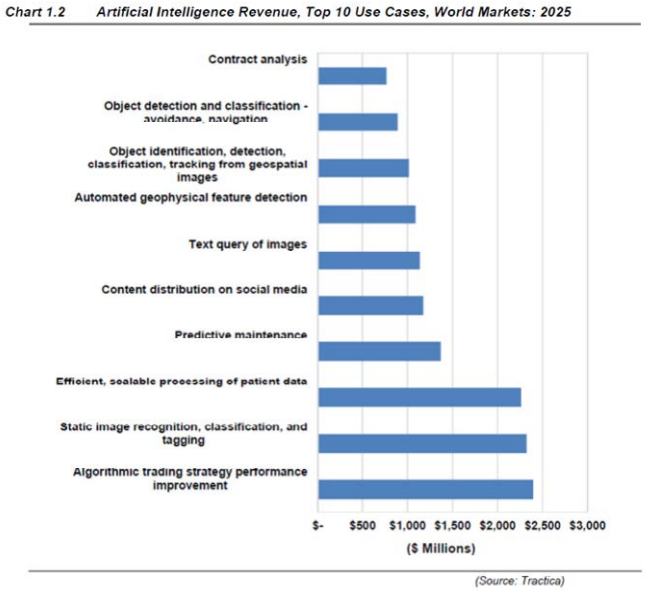

Fig.3 AI Revenue Top 10 Use Cases, World Market 2025[10]

CONCLUSION

GPT-3 is an autoregressive model that works on large corpus NLP datasets with strong performance. In Reinforcement learning human labels are used to train the models of reward and optimize model to provide feedback. Open AI system motivates many fine tune modeling applications like face recognition, text summarization, cloze activity task, sentence analysis. ROUGE is used for automatic summarization of text and to translate human text into machine readable language. There is huge scope for deep learning algorithms and GPT-3 to work on large corpus data for object detection, speech recognition, language translation and complex decisions with better result. Due to industrial revolution, smart phone technologies, IOT, cloud based applications there will be huge investment in artificial intelligent projects by 2025.

REFERENCES

1) https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-2ce44bb692ac
2) https://github.com/openai/gpt-3
3) https://en.wikipedia.org/wiki/Autoregressive_model
4) https://www.educative.io/blog/top-uses-gpt-3-deep-learning
5) Prafullahariwal, Heewoo Jun et.al.(Apr 2020), “Jukebox: A Generative Model for Music”,[Online] Available: https://openai.com/papers/, arXiv:2005.00341
6) Daniel M. Ziegler, Nisan Stiennon,et.al. (Jan 2020),“Fine-Tuning Language Models from Human Preferences”, [Online] Available: https://openai.com/papers/, arXiv: 1909.08593v2 [cs.CL]
7) Tom B. Brown, Benjamin Mann, et.al. (July 2020)” language Models are Few-Shot Learners”, [Online] Available:https://openai.com/papers/, arXiv: 2005.14165v4 [cs.CL]
8) Amanda Askell, Miles Brundage etal.,(July 2019) “The Role of Cooperation in Responsible AI Development” Source: https://openai.com/papers/, arXiv:1907.04534v1 [cs.CY]
9) Nisan Stiennon, Long Ouyang,et.al., (Oct 2020)“ Learning to summarize from human feedback”,Source:https://openai.com/paper, arXiv:2009.01325v2 [cs.CL]
10) 10.https://www.top500.org/news/market-for-artificial-intelligence-projected-to-hit-36-billion-by-2025/

xxxxxxxxxxx