Advances in Manufacturing Technology XXXV
M. Shafik and K. Case (Eds.)

doi:10.3233/ATDE220603

Research into Early Stage Identification of Entrepreneurs and Artificial Intelligence: Towards Self-Sustained Tools to Identifying Young Entrepreneurs

Stephen Robert PEARSON and Mahmoud SHAFIK
University of Derby, Kedleston Rd, Derby, DE22 1GB, UK.
College of Science & Engineering

Abstract: Youth unemployment is increasing and some countries are exhibiting unprecedented levels of youth unemployment, which according to research from Price Waterhouse Cooper (PwC), is costing the British economy £45bn per year, as well as blighting the careers of workers who miss out on a job in their teens and twenties. Unemployment exists because jobs do not, therefore one way to act to reduce it is to create jobs. It isn't the governments remit to create jobs, not in the private sector therefore this responsibility is falling more and more on entrepreneurs. This paper seeks to examine the potential use for Artificial Intelligence (AI) in the early-stage identification of young entrepreneurs, in order to provide a more accurate comprehensive survey tool, able to adapt to the questions and answers given during the identification process.

Keywords. Young Entrepreneurs, Youth Entrepreneurs, Entrepreneur Traits, Trait analysis, Artificial Intelligence, Entrepreneur Identification.

1. Introduction

Artificial intelligence (AI) is a rapidly evolving area where the Organisation for Economic Co-operation and Development (OECD) emphasises the importance of a collaborative approach to develop trustworthy AI systems for the benefit of society. Several work streams to measure AI-related trends are ongoing across the OECD. Established as an academic discipline in the 1950s, AI remained an area of limited interest for over half a century. Today, due to the rise of big data and improvements in computing power, it has entered the business environment and public policy worldwide. The last two decades have witnessed major advances in AI [7]. Future progress is expected to be even more influential, with many studies predicting that these technologies will transform work and business life around the world [2], [6], [11], [13], [14]. At the same time Brynjolfsson, Rock, and Syverson [3] shows that although systems using AI match or surpass human performance in more and more domains, it does not yet seem to have had a notable impact on productivity growth in the US. This paper seeks to examine the potential for the use of AI in the identification process of Young Entrepreneurs and whether its use can facilitate a more accurate and comprehensive survey tool for their identification.

2. Artificial Intelligence (AI) and Industrial Services Applications

Recent studies also highlight that an increasing number of economic models need to be updated in the light of AI-driven automation [1]. Today we are far from a satisfactory understanding of the impact of AI in the market, and statistically sound measures of AI diffusion and uptake will be essential to help update economic models that in turn can guide policy decisions. Equally, it is argued that AI will create new opportunities for business value creation and cost reduction [14]. The OECD definition of AI is; "An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy" [12]. While OECD provides an AI definition in its May 2019 Council Recommendation on Artificial Intelligence, when it comes to measurement and metrics, there is no official statistical definition of AI. A statistical AI definition provides information on what to measure and has a direct impact when implementing an AI module within ICT usage survey by businesses. For example, Eurostat, in its Community Survey on ICT Usage and E-commerce in Enterprises 2020, decided to use an "embedded approach", namely to include questions on AI in other relevant modules of the questionnaire such as big data, robotics etc., as the proposed definitions were found too complicated. It also considered that AI is not a standalone technology but co-exists and is embedded in other technologies.

Existing official measures are struggling to keep up with the rapid pace of AI. At the same time there is a pressing need to identify and share indicators to show where and how AI is developed, used, by whom, how fast, and in which sectors. To be able to deliver up-to-date policy relevant statistics it is suggested that official statistical information systems will need to build partnerships with businesses and academia. AI is a new and emerging area and it takes time for official statistics to incorporate it. In accordance with the official AI definition discussed earlier there are two areas that are problematic. The first one is the notion of a system with intelligence, i.e. the ability to think, which is vague. AI researchers use the notion of rationality to describe the term intelligence. It refers to the ability to choose the best action to take in order to achieve a certain goal, given certain criteria to be optimised and the available resources. The second problematic area is the term AI-system. It is difficult to capture as AI-systems usually comprise AI-based components embedded in a larger system, rather than a stand-alone system. Many times AI users in firms may not even know that they are using AI. That is why some AI definitions include examples. The McKinsey definition lists examples where machines have the ability to perform physical tasks using cognitive functions such as physical robotics, autonomous driving, and manufacturing work. Below are some examples of AI definitions that guides how the survey questions are designed. There seems to be a lean towards a common wording when describing what AI is. AI is transforming businesses and offers both possibilities and challenges. Finding the right balance between the two is not an easy task and further knowledge is required to capture all the nuances. AI is developing in a rapid pace, which means that it will continue to be somewhat of a moving measurement target.

3. AI and Social Media Resources

Artificial intelligence (AI) as previously explained, is the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. Machine learning gives computers the ability to learn without being explicitly programmed. Together, they could create a simplified, faster and more natural online survey interface. The survey of the future will give the impression of instant reaction, minimal wait time, and that the questionnaire is "learning" your preferences and predicting your responses. Think about the way a person interacts with Facebook and other social media. Each time they visit it is different, depending on their device, what they have posted, what others have posted, and what advertisers are targeting them. While the jury is still out about those "stalker ads" that seem to follow people from app to app, that sort of experience is the norm for the digital native. According to Felix Rios of Exomar "While the participant is answering each question, in the background the survey engine has to be loading the rest of the survey, prioritising the next question in line. Artificial intelligence should help us predict what is the most likely path this participant will take and prioritise those in the loading queue. Pretty much in the same way that Google starts to suggest search terms as we start typing a query in the search bar. These terms are highly relevant, and it's not by chance. The system can't wait until the participant clicks next to trigger the next steps. By the moment the participant is halfway through a 5×5 grid, the survey platform should have predicted what will happen next. This needs to happen in real time as he is interacting with the questions. Machine learning will make sure that the system becomes more and more accurate as it successfully repeats the prediction process. This is achieved using 'Repeated Incremental Pruning to Produce Error Reduction' (RIPPER) [19] that exist with the ability of rule generation. The decision tree is one of the most common rule-based classification algorithms among these techniques because it has several advantages, such as being easier to interpret; the ability to handle high-dimensional data; simplicity and speed; good accuracy; and the capability to produce rules for human clear and understandable classification [20] [21]. Online surveys today are modelled after paper surveys, a design that has gone by the wayside for younger respondents. Scrolling is the new preferred technique. As long as we continue to evolve the online survey to replicate the digital experience for respondents, the survey as a data collection method will remain viable. And AI will contribute to that.

4. AI as a Core Engine of Survey Tools

Completing online surveys can often feel like a never-ending onerous task, which to date has adopted a very 2-dimensional approach using the types of questions shown below.

- binary
- Multiple choice questions.
- Rating scale questions.
- Likert scale questions.
- Matrix questions.
- Dropdown questions.
- Open-ended questions.

- Demographic questions.
- Ranking questions.

Where AI does help is in creating an experience that is very comfortable for the respondent because it looks and feels like a chat session. The informality helps respondents feel more at ease and is well-suited to a mobile screen. The possible downside is that responses are less likely to be detailed because people may be typing with their thumbs.

4.1 Open-ended questions

There are three advantages to how AI treats open-ended questions. First, it can provide a first pass at codifying a thematic analysis of the data. When you go through the findings, the machine will have already grouped them according to the thematic analysis the AI has parsed. If you are using grounded theory (i.e., looking for insights as you go), this can be very helpful in getting momentum towards developing your insights.
Secondly, the AI also facilitates the thematic analysis by getting each respondent to help with the coding process themselves, *as part of the actual survey*. After the respondent answers "XYZ," the AI tells the respondent that other people had answered "ABC," and then asks if that is also similar to what the respondent meant. This process continues until the respondents have not only given their answers but have weighed in on the answers of the other respondents (or with pre-seeded responses you want to test). The net result for the researcher is a pre-coded sentiment analysis that you can work with immediately, without having to take hours to code them from scratch. The downside of this approach is that you will be combining both aided and unaided responses. This is useful if you need to get group consensus to generate insights, but it's not going to work if you need completely independent feedback. The third advantage of this approach over moderated qualitative methodologies is that the output can give you not only coded themes but also gauge their relative importance. This gives you a dimensional, psychographic view of the data, complete with levels of confidence that can be helpful when you look for hidden insights and opportunities to drive communication or design interventions.

4.2 Cross-platform and self-learning ability

There is an idea out there that AI surveys can access ever-greater sources of data for an ever-broader richness of insight. Yes, and no. Yes, we can get the AI to learn from large pools of respondent input. But, once again, without two-factor human input (from respondents themselves and the researcher), the results are not to be trusted because they run the likely danger of missing underlying meaning.

4.3 Creates real-time, instant surveys automatically

The final claim we need to address is that AI surveys can be created nearly-instantaneously or even automatically. There are some tools that generate survey questions on the fly, based on how the AI interprets responses. It's a risky proposition. It's one thing to let respondents engage with each other's input, but it's quite another to let them drive the actual questions you ask. An inexperienced researcher may substitute respondent-driven input for researcher insight. That said, if AI can take away some drudgery from the development of the instrument, as well as the back-end coding, so

much the better. “Trust but verify” is the way to go. So, this quote from Picasso may still hold true: “Computers are useless. They can only give you answers,” but now they can make finding the questions easier too. (Mark Szabo, 2019)

5. Conversational AI

Conversational artificial intelligence (AI) refers to technologies, like chatbots or virtual agents, which users can talk to. They use large volumes of data, machine learning, and natural language processing to help imitate human interactions, recognizing speech and text inputs and translating their meanings across various languages. Conversational AI combines natural language processing (NLP) with machine learning. These NLP processes flow into a constant feedback loop with machine learning processes to continuously improve the AI algorithms. Conversational AI has principal components that allow it to process, understand, and generate response in a natural way. Machine Learning (ML) is a sub-field of artificial intelligence, made up of a set of algorithms, features, and data sets that continuously improve themselves with experience. As the input grows, the AI platform machine gets better at recognizing patterns and uses it to make predictions. Natural language processing is the current method of analysing language with the help of machine learning used in conversational AI. Before machine learning, the evolution of language processing methodologies went from linguistics to computational linguistics to statistical natural language processing. In the future, deep learning will advance the natural language processing capabilities of conversational AI even further. NLP consists of four steps: Input generation, input analysis, output generation, and reinforcement learning. Unstructured data transformed into a format that can be read by a computer, which is then analysed to generate an appropriate response. Underlying ML algorithms improve response quality over time as it learns [19][20][21]. These four NLP steps can be broken down further below:

- Input generation: Users provide input through a website or an app; the format of the input can either be voice or text.
- Input analysis: If the input is text-based, the conversational AI solution app will use natural language understanding (NLU) to decipher the meaning of the input and derive its intention. However, if the input is speech-based, it’ll leverage a combination of automatic speech recognition (ASR) and NLU to analyse the data.
- Dialogue management: During this stage, Natural Language Generation (NLG), a component of NLP, formulates a response
- Reinforcement learning: Finally, machine learning algorithms refine responses over time to ensure accuracy

When people think of conversational artificial intelligence, online chatbots and voice assistants frequently come to mind for their customer support services and omni-channel deployment. Most conversational AI apps have extensive analytics built into the backend program, helping ensure human-like conversational experiences. Experts consider conversational AI's current applications weak AI, as they are focused on performing a very narrow field of tasks. Strong AI, which is still a theoretical concept, focuses on a human-like consciousness that can solve various tasks and solve a broad range of problems. Despite its narrow focus, conversation AI is an extremely lucrative technology for enterprises, helping businesses more profitable. While an AI chatbot is

the most popular form of conversational AI, there are still many other use cases across the enterprise. While most AI chatbots and apps currently have rudimentary problem-solving skills, they can reduce time and improve cost efficiency on repetitive customer support interactions, freeing up personnel resources to focus on more involved customer interactions. Overall, conversational AI apps have been able to replicate human conversational experiences well, leading to higher rates of customer satisfaction. Conversational AI is still in its infancy, and while the idea of interacting with a computer using voice or text goes back a long way, it is only in recent years that this idea has become a reality with the emergence of digital personal assistants, smart speakers, and chatbots. Advances in AI, particularly in deep learning, along with the availability of massive computing power and vast amounts of data, have led to a new generation of dialogue systems and conversational interfaces, resulting in widespread business adoption in recent years [10]. As with any new technological advances, there are some challenges with transitioning to conversational AI applications.

5.1 Language input

Language input can be a pain point for conversational AI, whether the input is text or voice. Dialects, accents, and background noises can impact the AI's understanding of the raw input. Slang and unscripted language can also generate problems with processing the input. However, the biggest challenge for conversational AI is the human factor in language input. Emotions, tone, and sarcasm make it difficult for conversational AI to interpret the intended user meaning and respond appropriately.

6. Chatbot use for the identification of Young Entrepreneurs

By selecting the right questions as part of the selection process, coupled with a comprehensive range of effective answers, the use of chatbots to engage with the potential entrepreneurs to help extract information and responses from them that may otherwise not have been readily accessible by use of simple Yes/No or Likert Scale responses. Chatbots coupled with a level of AI that would enable them to interact with respondents has the potential not only to engage with them in a stimulated way, but also ask pertinent questions, the answers to which can be immediately analysed and further utilised to ask more questions and create new ones relevant to the responses given. This sort of technology is the way forward for interactive extraction of information in an evaluation of a candidates' suitability for selection as a potential entrepreneur and should be further explored in studies following this one.

7. Conclusions

Artificial intelligence (AI), particularly, *machine learning (ML)* have grown rapidly in recent years in the context of data analysis and computing that typically allows the applications to function in an intelligent manner [14]. ML usually provides systems with the ability to learn and enhance from experience automatically without being specifically programmed and is generally referred to as the most popular latest technologies in the fourth industrial revolution (4*IR* or Industry 4.0) [15] [16]. It can therefore be concluded

that based upon the developments of AI, the use of conversational AI in the identification of potential entrepreneurs could be a useful enhancement to aid in a successful identification process. As we are dealing with young people who are used to social media, interaction via chat type apps with their peers, then the use of 'Chatbots' and associated media would serve to extract the relevant information from them, putting them at ease with the process in what would be a familiar environment to them. As AI improves and advancements are made, the survey process could become almost like an interview with a fellow human, with the AI being able to assess and determine what the next question should be, based upon the previous answers, and try to create a seamless identification process. This paper is part of ongoing research programme, and the future findings will be reported as soon as the data is available and fully analysed. It is expected that the outcomes of this studies will be combined with the VR tools to develop a long-term educational tools for potential young entrepreneurs.

References

[1] Acemoglu, D., & Restrepo, P. (2018). Artificial Intelligence, Automation and Work. MIT working paper.

[2] Brock, J. K.-U., & von Wangenheim, F. (2019). Demystifying AI: What Digital Transformation Leaders Can Teach You about Realistic Artificial Intelligence.California Management Review, 61(4), 110-134.,Doi:10.1177/1536504219865226

[3] Brynjolfsson, E., Rock, D., & Syverson, C. (2017). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. NBER.

[4] DIGG. (2019). Promote the public sectors' ability to use AI. Retrieved from EC. (2019). A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES.

[5] Deepa Mani, Madhumitha Mohan, and Monish K, "The AI Index 2021 Annual Report", Srini Raju Centre for IT and the Networked Economy, Indian School of Business, May 2021

[6] Garbuio, M., & Lin, N. (2019). Artificial Intelligence as a Growth Engine for Health CareStartups: EMERGING BUSINESS MODELS. California Management Review, 61(2), 59-83. doi:10.1177/0008125618811931

[7] Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. California Management Review, 61(4), 5-14. doi:10.1177/0008125619864925

[8] House of Commons Library (2017). Youth Unemployment Statistics.

[9] Huq, A., & Gilbert, D. (2017). All the world's stage: transforming entrepreneurship education through design thinking. *Education+ Training*, *59*(2), 155-170.

[10] Mattias Wahde, Marco Virgolin. (2022) DAISY: An Implementation of Five Core Principles for Transparent and Accountable Conversational AI. International Journal of Human–Computer Interaction 11, 1-18.

[11] McKinsey. (2019). Global AI Survey: AI proves its worth, but few scale impact. Retrieved from McKinsey & Company:

[12] MGI. (2017). A Future That Works: Automation, Employment, and Productivity. Retrieved from OECD. (2019a). Artificial Intelligence in Society. Paris. OECD. (2019b). Proposal for the OECD AI Policy Observatory.

[13] PWC. (2017). Sizing the prize: What's the real value of AI for your business and how can you capitalise? Retrieved from PWC.(2018). The macroeconomic impact of artificial intelligence.

[14] Ransbotham, S., Shervin, K., Fehling, R., LaFountain, B., & Kiron.(2019)

[15] Sarker IH. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Comput Sci.

[16] Sarker IH, Hoque MM, MdK Uddin, Tawfeeq A. (2020). Mobile data science and intelligent apps: concepts, ai-based modeling and research directions. Mob Netw Appl, 1–19.

[17] Sarker IH, Kayes ASM, Badsha S, Alqahtani H, Watters P, Ng A. (2020). Cybersecurity data science: an overview from machine learning perspective. J Big Data.;7(1), 1–29

[18] Spillman, L. (2014). Mixed Methods and the Logic of Qualitative Inference. Qualitative Sociology, 37(2), 189-205. doi:10.1007/s11133-014-9273-0

[19] Taatila, V. P. (2010). Learning entrepreneurship in higher education. *Education+ Training*, *52*(1), 48-61.
[20] Witten IH, Frank E, Trigg LE, Hall MA, Holmes G, Cunningham SJ. Weka (1999). Practical Machine Learning Tools and Techniques with Java Implementations.
[21] Wu C-C, Yen-Liang C, Yi-Hung L, Xiang-Yu Y. (2016). Decision tree induction with a constrained number of leaf nodes. Appl Intell.; 45(3):673–85.
[22] Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Philip SY, et al. (2008). Top 10 algorithms in data mining. Knowl Inform Syst;14(1):1–37