



The Impact of Machine Learning on the Future of Insurance Industry

Harish Paruchuri

Senior AI Engineer, Department of Information Technology, Anthem, Inc., USA

E-mail for correspondence: harishparuchuri9999@gmail.com

ABSTRACT

Recently data remains the central and the core concentration in the insurance industry. The outburst in data generation thus created the need for technologies that will be used to process or manage big data in the industry. Thus, the conditions at hand are sighted by a plodding but firm change, which is compelled by an atmosphere shown by enhanced rivalry, fraud activities, flexible market places, high prospects from clients, and stringent guidelines. The introduction of machine learning in solving industry tasks in the assurance value chain such as underwriting and forfeiture avoidance, entitlements management, fraud uncovering, product evaluating, transactions, and client capability will put the industry in the damp light in the future due to high increase of big data. This paper has examines some cases and bring out the vital role of machine learning in handling client data and resolving issues of entitlements. Hence, machine learning holds a brighter future for insurance organizations if implemented well.

Keywords: Machine learning, insurance industry, data, assurance value chain

INTRODUCTION

The most essential benefit of machine learning (ML) in the insurance industry is to simplify data sets. Thus ML can successfully simplify datasets in organized, unorganized, and semi-organized datasets (Jones et al., 2019). It accurately provides a value chain in identifying risk, client's action, and entitlement using innovative forecasting exactness. The actual uses of machine learning in the insurance industry are enormous, they range from insightful threat enthusiasm, premium outflow to outflow running, subrogation, accounts, and scam uncovering (Burri et al., 2019). Machine learning has been in existence for a while now and the most common classes of machine learning include supervised learning, unsupervised learning, and reinforcement learning (Belhadji et al., 2012). So far many researchers have identified many stakeholders employing controlled learning for the last 10 years in assessing the risk using the known variable in diverse arrangements to achieve their goal(s) (Crocker and Tennyson, 2010). With the current trends, stakeholders have started exploring unsupervised learning, with it, predetermined objectives are rich (Kajiamuller, 2013). When there are variations in the parameters, the approach

detects those variations and attempts to adjust as per the objectives. For instance, based on traffic, the global positioning system recommends several ways or directions vigorously according to traffic situations (Burri et al., 2019). Reinforcement learning depends on the artificial neuron network (ANN), goa/.target can be altered vigorously based on the objective. Hence, corroboration learning makes use of internet of things (IoT) applications (Kotsiantis, 2011). These types of machine learning, the ones explored so far and those yet to explore conveys a positive future for the insurance industry.

Statement of Problem

Big data is the main focal point in the insurance industry, and as of now, insurance haulers have the right of entry to more insurance data than ever in the past. More data have been generated in the last 24 months than the human race has ever witnessed. Many stakeholders in the industry have been flabbergasted by the flare-up in the data generated from a host. These sets of data include social media movement, telematics, voice analytics, online, wearable devices, and networking sensors (<https://www.forbes.com/forbes>). This outburst in data generation thus created the need for

technologies that will be used to process or manage big data in the industry. Thus, the conditions at hand are sighted by a plodding but firm change, which is compelled by an atmosphere shown by enhanced rivalry, fraud activities, flexible market places, high prospects from clients, and stringent guidelines. Stakeholders are being compelled to explore possible ways to be employed in forecasting demonstration and ML to manage their rivalry edge, improve industry process and improve client's happiness (Burri et al., 2019; <http://www.propertycasualty360.com>). Also, insurers are examining how they can take advantage of the evolution of technological approaches like machine learning in solving industry tasks in the assurance value chain such as underwriting and forfeiture avoidance, entitlements management, fraud uncovering, product evaluating, transactions, and client capability (Vadlamudi, 2016).

Objectives of this study

This study focuses on the impact of machine learning on the future of the insurance industry. We will explain how insurance data can be unlocked, the prospective for ML in the industry, and challenges that may hinder the implementation of machine learning.

LITERATURE REVIEW

Machine Learning

ML is the art of offering computers the capacity to learn from data devoid of being overtly programmed (Geron, 2019). It's a technique of data analysis that has been improving high requests over the last eras, as it empowers individuals and firms the capability to appreciate their datasets in a broader view and with details. According to Forbes investigation proposes that one out of ten organizations or more are using machine learning. Most of them use it to ascertain the effectiveness of scam detecting, process optimization, and opinion mining (Vadlamudi, 2019). Machine learning is an extension of artificial intelligence and from the application of logic and situations. It enables technology to enhance at a particular job with capability, by learning from the data and detecting models of importance with nominal human interference (SAS, 2020).

Machine learning deals with big data, which is gradually everywhere, and they are widely referred to as a source for an improved understanding economy. More so, for a better understanding of big data in insurance may be possibly employed big data as its source of data analytics. Vadlamudi (2015) defines big data with 3Vs which is associated with 3 problems in data management; variety, volume, and velocity (Chen et al., 2012), which are all high in the subject of big data. High volume is the proportions of the data that is represented in multiple terabytes. High variety is related to operational heterogeneity in the dataset that institutes about 10 – 15% of all present data, denotes to the horizontal data originate in a worksheet or interactive databanks. Illustrations of unstructured data include images, text, video, and audio, which at times lack

the operational arrangement needed by the mechanisms to be applied for breakdown and varieties up to about 80% or more of all firm data. Traversing a band between entirely organized and shapeless data, the format of semi-organized data prepares not to adapt to severe criteria. Software for text exchange on the web like extensive markup language is a common example of semi-organized data composes about 5-10% of data (Gandomi and Haider, 2011; Taylor, 2018).

Velocity signifies the rate at which data is generated and how rapid is it to analyze or act upon. The spread of alphanumeric tools like sensors and smartphones has paved the way to a unique speed of data that is created (Gandomi and Haider, 2015). On the other hand, it entails a high rate at which data is required to strain. FCA (2016) investigated an input report outlining the actual meaning of big data. This literature discussed big data as;

- Utilizing novel or prolonged datasets and data, covering from eccentric bases including social media;
- Implementing the technologies essential to produce, gather and save these new methods of data;
- Utilizing advanced data treatment methods;
- Refined analytical methods such as predictive analytics; and
- Applying this data understanding in professional verdicts and events.

Hence, this approach is not restricted to only one set of data, but also comprises data analytic and processing features of big data. Furthermore, FCA literature pinpoints the primary origin of big data as:

- Exclusive data (e.g., data from associated firms such as individual data of products bought or loyalty vouchers)
- Data picked up from secondary sources (example includes combined search engine data like credit payments, license particulars, entitlements concession databases, bill appraisal website costing)
- Social media data (example includes clients-explicit data obtained from Twitter or Facebook) and
- Linked tool data (example includes telematics equipment that could be applied in the home, motor, or health telematics).

Exclusive and acquired data all together linked device data would possibly be organized data, together with data generated from IoT. Data generated from social media may be possibly unorganized that makes it expensive to be used for big data analytics goals. Organized data can also contain visual and audio data that would be of assistance during tragedies. Artificial intelligence is strictly associated with big data but different. Artificial intelligence is the approach of machine learning that applies data to realize the learning approach of the algorithms. Artificial intelligence is an explicit method that may or may not apply big data, whereas big data is the data that may be entered for a variety of procedures in the insurance value chain.

Where and how Machine Learning may affect Insurance Industry

Machine learning is predicted to impact the insurance industry in diverse ways, but the most estimated is the data analytics, underwriting, and assessing with diverse opinions in what way machine learning could impact them. Supplies and transactions are more understandable means owing to the way machine learning might improve targeting and appreciation of user behavior. Entitlements management and grievances may be rationalized utilizing machine learning, and promoting may be more targeted with big data. Data analytics is the art of illustration understandings from uncooked information bases. It is a comprehensive subject that involves several diverse kinds of data testing. Any kind of data may be exposed to data analytic methods to receive an understanding that might be applied to enhance insight and procedures. Several methods and procedures of data analytics are being computerized into power-driven procedures and algorithms.

Essentially, data analytics are classified into 4 classes (Investopia, 2018):

- *Descriptive analytics*: This involves the analysis of data over some time.
- *Diagnostic analytics*: this involves many different data inputs and some postulating.
- *Predictive analytics*: this deals with the future and what might happens.
- *Prescriptive analytics*: this class of data analytic deals with a sequence of activities going onward.

According to data analytic devices, stakeholders may take pride in machine learning to relate problem-solving and extrapolative analytics to estimate the performance of likely policyholders as well as acting according to the outputs (OECD, 2020).

It is noteworthy to remark that devoid of skilled input, machine learning analytics may be focused on forged relation, and attention requires to be reserved in extrapolating data. Forged relations exist when 2 arbitrary parameters trail one another meticulously on a graph, prominent one to dubious relations. This might lead to expectations that one of the parameters is connected to the measure of the other parameters (Vadlamudi, 2018). On the other hand, this statement does not portray a relationship among the parameters, but only from the skilled investigation will the relationship can be well-known. Among firms in the insurance industry, they essentially depend on conventional generalized linear models (GLMs) to examine price risk (Vadlamudi, 2016). This procedure is a statistical approach apply to predict the connections among the possibility of creating entitlement and diverse risk variables. Moreover, many organizations have commenced the application of analytical methods all together to make inputs to generalized linear models. For instance, decision tree analytics or non-linear procedures like machine learning methods are being sorted after (FCA, 2016).

The GLM is an overview of the general linear model. To simplify it, a linear model stipulates the linear connection among a dependent parameter, Y, and a regular of forecaster parameters, the X's, so that $Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$. In this equation, b_0 is the regression coefficient for the intercept and the b_i standards are the regression coefficients (for variables 1 through k) calculated from the data.

The International Association of Insurance Supervisors (IAIS) is aiming through the production lifespan of insurance, and discerning on some cases when machine learning is considered. Machine learning is being engaged in product growth, for instance, via the application of telematics. Whereas most extensively identified and essentially functional, practice case is auto telematics, and a good number of telematics are there that will be incorporated into insurance products in near future.

Major Factors motivate the use of Machine learning in insurance

According to Accenture (2018) stresses the major factor that motivates the use of machine learning in the insurance industry as follows:

- i. Smart everything – firms are looking at possible ways to harness the advanced ML to energize smart, or automation in several sectors like healthcare diagnosis, estimation maintenance.
- ii. As data keep growing, open-source models like machine learning are helping in the management of big data in the insurance industry.
- iii. The data size and velocity pave way for the emergence of machine learning in monitoring and working out instead of programming structure becoming important.
- iv. Machine learning helps in the ability to talk back and facial respect, serving to create it more valuable.

Figure 1 demonstrates the growth of the machine learning marketplace in different continents over a period of ten years. It displays the hastening acceptance of machine learning and the serious significance of this trending technology.

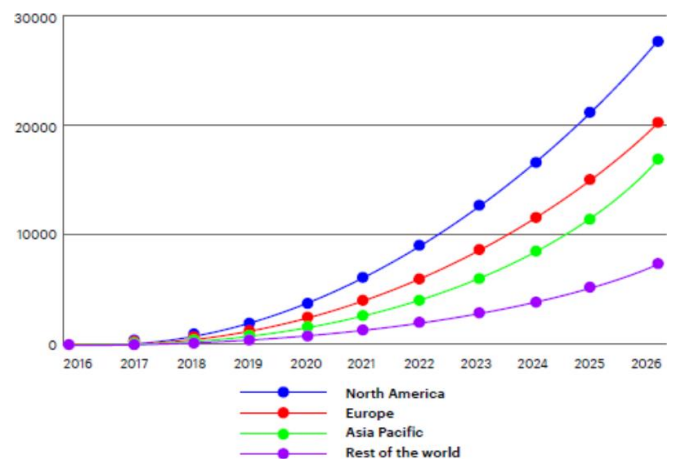


Figure 1: Global Machine learning market by continents between 2017-2026 (in United States Dollars million)

Some potential role machine learning will play in the insurance value chain cannot be overemphasized. These range from data management to initial interaction to determine which one covers a client requirement. However, machine learning is extensively apply across the assurance value chain as shown in Figure 2 (Accenture, 2018).

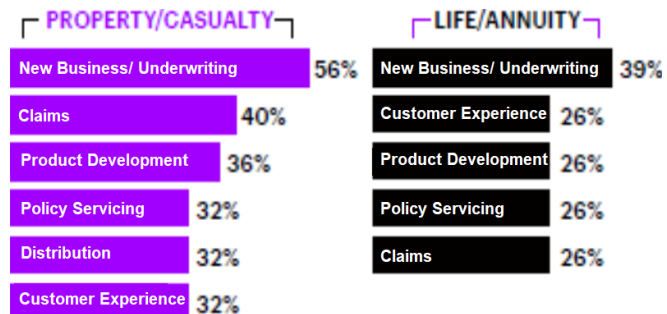


Figure 2: Potential of machine learning in the assurance value chain

METHODS

To achieve the main goal of this study, which is to assess the impact of machine learning on the future of the insurance industry. Also, provide a competent explanation on how insurance data can be unlocked, forthcoming for ML in the industry, and challenges that may hinder the implementation of machine learning. We adopted some cases for machine learning in the insurance industry and offered a further description of how it will be instigated. It is interesting to know that all cases used in this study will enhance existing competencies, somewhat than replace human flow, even though many are associated with automation.

Research Questions

Hence, the following research questions were attempted to be answered using different cases;

- What can we study about ML's potential impact on the future of the insurance industry using cases?
- What are the challenges hindering the implementation of machine learning in the insurance industry?

RESULTS AND DISCUSSION

Inspection of Cases Used

Among the commercial companies, the insurance industry performs an inspection to ascertain their underwriting decisions according to the information gathered from the hazard; pre-cover, post-initiation with the rekindling cycle. This helps them to recognize any occurring or possible dangers and assist their clients in disaster management, in that way, decreasing their exposure. As a result of the complication and/or volume of the occupation, although inspections are time-consuming to some extent. The cases

employed in this study aims at the elimination of the inspection period and improve surveyor productivity with the use of machine learning (OECD, 2020).

The application of machine learning will create deep learning image appreciation patterns in the insurance industry and work out utilizing labeling and augmenting datasets to classify hazards. Working out the pattern on a larger scale and augmenting datasets will aid in improving the model precision and eliminate the danger from data favoritism that would have factored in with little data sets. Immediately the model is worked out, it can be transformed to core machine learning for application with a portable App engraved with swift utilizing outlines such as vision to deduce core machine learning patterns with device identification abilities. With this surveyor with capability is enhanced to process the recorded drone footage and assess on the spot. It can also be achieved the other way round when the recorded drone footage is collected and upload for processing using the cloud podium as a consignment career in the office.

Introduction of these objects, both automated or surveyor supervised, or body-cameras to perform inspections help in reducing the inspection time. Outcomes can be generated utilizing parsing results of the activities of the device detective models to classify by severity and assigned a combining machine learning with the inspection yields richer data for the insurance industry. This approach will not only be profitable to the surveyor while gathering the inspection account but could assist the underwriter in taking well-versed underwriting choices and eventually aid customers with their disaster management.

Internet of Things Case

One of the popular prospects of machine learning in the insurance sector is with the IoT. Insurance firms are progressively affiliating with the insurance-based technology in order to increase their client's shares by taking merits of groundbreaking results (Sivarajah et al., 2017).

Also, machine learning provided the insurance sectors with collaboration with the technological sector to help them develop an application that will help in solving some of their greatest challenges in the sector. Like the InsurTechn is an open prospect brought to the strength of the sector by machine learning approaches. This application will insurers to provide discounted packages to their clients according to the InsurTech designers (BaFin, 2017).

There are many internets of things smart home objects deployed in the insurance industry that have the ability to detect and alerting the client if there is any problem within or in their home or commercial property, for instance, moisture or leak sensor. By coalescing the internet of things and machine learning in the insurance industry can provide better services to their clients (Anyoha, 2017).

In the future if the insurance industry can leverage the data generated by the smart home objects, and locations even weather datasets they could use them to estimate and carry out analysis that will be of assistance to their clients apart from identifying but avert activities that are disastrous. For instance, when a customer had mounted leak identification objects in their property, estimated analytics protocols can be built or construct using the above-mentioned datasets to forecast which clients may be susceptible to leak. Then proactive action can be taken by stakeholders in the sectors by sending out renovators to substitute damaged tubes before they eruption leading to claims.

With expert python designers, these protocols can be constructed manually by deploying a Jupyter notebook and ML libraries like Scikit-learn. Alternatively, if there was no needed knowledge available in the house, this can be produced and distributed by deploying a data science workbench such as IBM Watson studio. According to Jones (2019), the deployment of the internet of things and machine learning as examined in different geographical zone shows an increasing rise (see Figure 3).

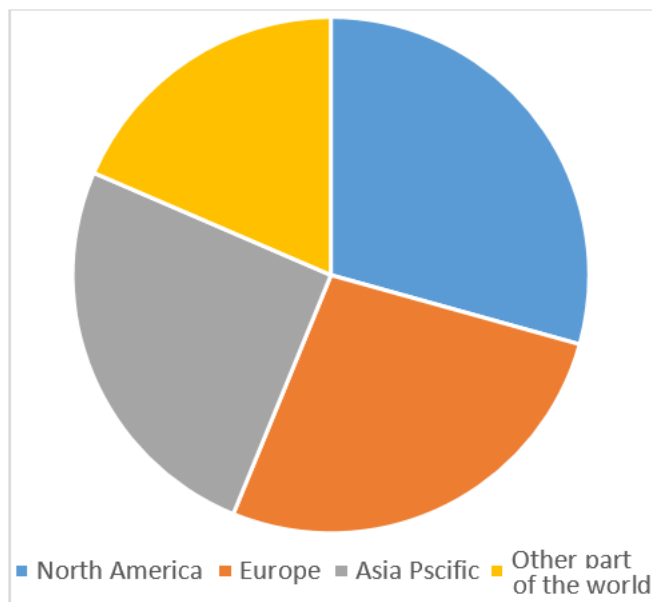


Figure 3: Contribution of Machine learning based on geographical zone

Entitlements Fast-tracked case

Machine learning can be used to quicken and automate entitlements. This has to do with way customer interacts with the insurer can be quickened with the innovation of machine learning. If it is shared with visual-oriented artificial intelligence image classification and machine learning to forecast the price and effect of the entitlements, a decision can be fast-tracked with full artificial intelligence oriented computerization leading to entitlements being paid in a few hours or days instead of weeks or months (OECD, 2019).

It seems like this type of computerization for entitlements fast-tracked will work only in fewer impact entitlements,

but the unfluctuating impact of applying artificial intelligence to manage these entitlements kinds can lead to more labor-intensive and procedure determined complex entitlements being increased with some of the artificial intelligence approaches applied in the event could free up agents to prioritize occupation and devote more time working on the complex entitlements.

It very obvious that entitlements such as healthcare insurance entitlements can be fast-track deploying artificial intelligence method to recap forms either handwritten to or applying natural language indication can be analyzed faster to make well-versed and precise decisions.

Some Challenges encounter in implementing Machine Learning

Below are some of the challenges stakeholders generally face when trying to implement machine learning:

1. *Training Necessities*

Artificial intelligence-oriented highbrow structures much are trained in a sphere, for example, entitlements or promoting for an underwriter. This demands a discrete training structure, which underwriters find difficult to offer for training the artificial intelligence protocols. Protocols are required to be trained with huge volumes of documents or transactions to cover all potential situations (Mishr, 2016).

2. *The right source of data*

The data quality deployed to work-out forecasting protocols is likewise essential as the data quality, especially when it is to with machine learning. The datasets required to be symbolic and stable so can take a quality picture and escape partiality. The aforementioned challenge is an essential factor when working out forecasting protocols. Usually, stakeholders are struggling to offer pertinent data for working out artificial intelligence protocols (Accenture, 2018).

3. *Exertion in forecasting returns*

It is not that easy to forecast enhancement that machine learning can usually allow to a scheme. For instance, it is not always an easy task to budget or plans a scheme using machine learning, as the funding required may differ per project, according to the investigations. Hence, it is utmost, not feasible to forecast the return on asset. This creates it hard to get everyone on the panel the idea and participate in it (Burri et al., 2019).

4. *Data security*

The enormous quantity of data castoff for machine learning algorithms has shaped an extra security risk for insurance firms. With such an upsurge in data collected and networking between bids, there is a threat of data leaks as well as security openings. A security event could pave the way to personal material dwindling into the incorrect hands. This generates anxiety in the minds of underwriters (Burri et al., 2019).

CONCLUSIONS AND RECOMMENDATIONS

The emergence of recent technologies is moving tremendously fast making their techniques into organizations. To this effect, the insurance sectors do not exhibit deficiency behind the others. The deployment of statistical tools in the insurance industry has a lasting record. Hence, the issue that insurance firms are actively applying the art of data analytics is not astonishing. In essence, the purpose of using the art of data analytics in the sector is the same as in the other sectors to improve marketing approaches, we offered many machine learning approaches to aid in the management and analyses of big data generated in the sector, thereby helping the stakeholders to analyses entitlements proficiently and liken their performance by means of different matrices. Therefore, we recommend machine learning to as many that are yet to implement this science of data management as it will help the stakeholders to manage their clients better in the near future.

REFERENCES

- Anyoha, R. (2017), The History of Artificial Intelligence, Harvard Science in the News, <http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/> (accessed on 23/04/2021).
- BaFin (2018), Big Data meets artificial intelligence– Challenges and implications for the supervision and regulation of financial services, https://www.bafin.de/SharedDocs/Downloads/EN/dl_b_dai_studie_en.html (accessed on 26/04/2021).
- Belhadji, E., G. Dionne, and Tarkhani, F. 2012. A Model for the Detection of Insurance Fraud, Geneva Papers on Risk and Insurance Theory, 25: 517-538, may 2012.
- Burri, R.D., Burri, R., Bojja, R.R. and Buruga, S.R. 2019. Insurance Claim Analysis Using Machine Learning Algorithms. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(6&4): 577-582.
- Chen, Chiang and Storey 2012. "Business Intelligence and Analytics: From Big Data to Big Impact", *MIS Quarterly*, 36(4): 1165, <http://dx.doi.org/10.2307/41703503>.
- Crocker, K. J., and Tennyson, S. 2010. Insurance Fraud and Optimal Claims Settlement Strategies: An Empirical Investigation of Liability Insurance Settlements. The Journal of Law and Economics, 45(2), April 2010.
- FCA. 2016, "FS16/5: Call for Inputs on Big Data in retail general insurance | FCA", No. FS16/5, Accessed on 25/04/2021 <https://www.fca.org.uk/publications/feedback-statements/fs16-5-call-inputs-big-data-retailgeneral-insurance>
- Gandomi, A. and Haider, M. 2015. "Beyond the hype: Big data concepts, methods, and analytics", *International Journal of Information Management*, 35(2): 137-144, <http://dx.doi.org/10.1016/j.IJINFORMGT.2014.10.007>.
- Geron, A. 2019. Hands-on machine learning with Scikit-Learn, Keras, and Tensor Flow: concepts, tools, and techniques to build intelligent systems. Accessed on 25/04/2020.
- Jones, S., Humphreys, B, and Woolnough, M. 2019. Considering the Impact of AI in Insurance. *IBM Power Systems*, 1-25.
- Kajiamuller, 2013. The Identification of Insurance Fraud – an Empirical Analysis Working papers on Risk Management and Insurance no: 137, June 2013.
- Kotsiantis, S.B. 2011. Supervised Machine Learning: A Review of Classification Techniques, *Informativacovl.*, 31, 249-268.
- Mishr K. 2016. Fundamentals of life insurance theories and applications. In: 2nd ed, Delhi: PHI Learning Pvt Ltd
- OECD (2020), The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector, www.oecd.org/finance/Impact-Big-Data-AI-in-the-Insurance-Sector.htm.
- SAS. 2020. Machine learning, what it is and why it matters. Accessed on 20/04/2021 from https://www.sas.com/en_gb/insights/analytics/machine-learning.html
- Sivarajah U, Kamal M, Irani Z and Weerakkody V. 2017. Critical analysis of big data challenges and analytical methods. *J Bus Res* 70:263–286
- The Accenture website <https://www.accenture.com>. Accessed on 24/04/2021
- Vadlamudi, S. (2015). Enabling Trustworthiness in Artificial Intelligence - A Detailed Discussion. *Engineering International*, 3(2), 105-114. <https://doi.org/10.18034/ei.v3i2.519>
- Vadlamudi, S. (2016). What Impact does Internet of Things have on Project Management in Project based Firms?. *Asian Business Review*, 6(3), 179-186. <https://doi.org/10.18034/abr.v6i3.520>
- Vadlamudi, S. (2018). Agri-Food System and Artificial Intelligence: Reconsidering Imperishability. *Asian Journal of Applied Science and Engineering*, 7(1), 33-42. Retrieved from <https://journals.abc.us.org/index.php/ajase/article/view/1192>
- Vadlamudi, S. (2019). How Artificial Intelligence Improves Agricultural Productivity and Sustainability: A Global Thematic Analysis. *Asia Pacific Journal of Energy and Environment*, 6(2), 91-100. <https://doi.org/10.18034/apjee.v6i2.542>

--0--