Evaluating the Use of Machine Learning Algorithms in Predicting Drug-Drug Interactions and Adverse Events during the Drug Development Process

Sri Sai Subramanyam Challa
Independent Researcher, USA.

Abhip Dilip Chawda.
Independent Researcher, USA.

Abhishek Pandurang Benke
Independent Researcher, USA.

Mitul Tilala
Independent Researcher, USA.

Abstract

This paper aims at reviewing the use of ML algorithms in the identification of DDIs and AEs during drug development. In the past, methods involved did not significantly allow for forecasting DDIs and AEs, largely due to historical data and lack of deeper understanding of pharmacology of drugs. ML algorithms are more efficient due to big data from clinical trials and patients' EHRs, finding complex relationships that enhance precision. In the paper, it looks at different forms of ML approaches such as regression and neural networks and clustering where a better performance is observed as compared with the conventional approaches. It can be concluded that, the proposed method of using ML for classification demonstrates a higher degree of accuracy, precision and recall compared with the other models and could be impactful in enhancing drug safety and efficacy in future. However, there is still issues such as data quality issues as well as ethical issues managing patient's data. Future development includes the use of different data and incorporating modern technologies into predicting ability and ethical issues.

Keywords: Machine Learning**,** Drug-Drug Interactions**,** Adverse Events**,** Pharmacovigilance.

DOI Number: 10.48047/NQ.2022.20.20.NQ109009 NeuroQuantology2022;20(20): 63-74

TABLE OF CONTENTS

Introduction

DDI and AEs risk prediction is crucial in the course of drug development with an aim of optimizing the therapeutic outcomes for patients. Historically, such tasks could not be solved with traditional approaches because of the sheer amount and density of pharmacological information. Machine learning algorithms are a more effective solution because these algorithms learn from large amounts of data to find patterns and estimate possible DDIs and AEs. The purpose of this research will be to review the use of diverse ML strategies in the field of pharmacovigilance in order to demonstrate how this area could be transformed in the future. This redundancy, aided by superior predictive models, can be further refined to fortify the pharmaceutical's industry's aptitude for risk identification and prevention, thus making for better medications on the whole.

Literature Review

Analysis of Adverse Events Database

According to Poleksic and Xie, 2019, they created an extensive library on adverse effects most closely related to the drugs and drug mixtures. In their study published in the journal Scientific Reports, the authors offer a useful tool for scientists and doctors who strive to decode the interactions that occur within patients' bodies. It holds all forms of information about different types of side effects related to individual medications and their interactions, which helps in analyzing the patterns of adverse effects in pharmacological interventions. It is for this reason that their database becomes a good reference in the usage of the machine learning applications in the pharmacovigilance. Rich and detailed enough to power the training and testing of machine learning algorithms that can be used to predict possible unwanted occurrences. Therefore, this study emphasizes on the need for proper data acquisition methodology and its incorporation into modern complex analysis models, which helps to improve the credibility of risk analysis of adverse effects of drugs and the overall advancement of accurate therapeutic approaches.

Statistical Methodologies for Detecting Drug-Drug Interactions

According to Noguchi *et al.*, 2019, it contains a thorough discussion and critical evaluation of statistical methods applied to identify DDIs through the SR systems,

according to the article in Frontiers in Pharmacology. Their review and critique describe the advantages and drawbacks of the main techniques in disproportionality analysis, Bayesian inference, and machine learning. These methodologies are very important in determining possible DDIs from large datasets generated from actual clinical environment.

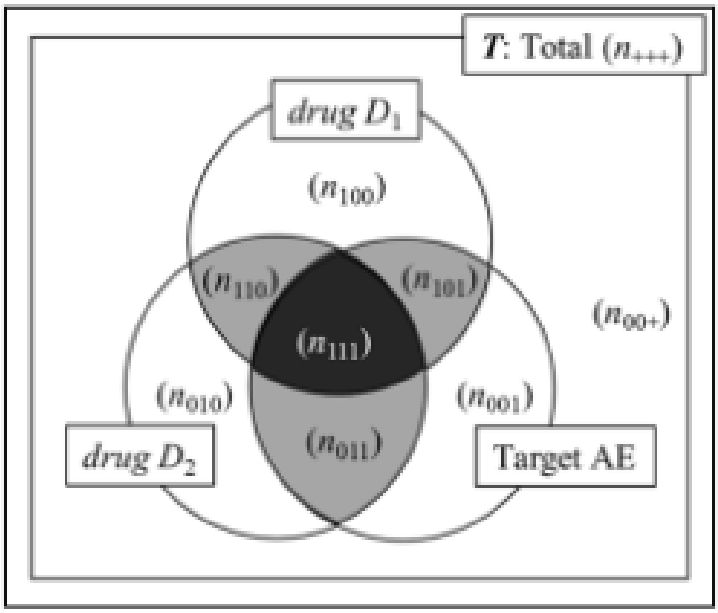

Figure 1: Venn diagram for the evaluation of drug–drug interaction
(Source: Noguchi *et al.,* 2019)

The authors stress the urgent need for effective signal detection to avoid risks and enhance patient outcomes. They elaborate that applying more sophisticated statistical analysis can improve the accuracy of the DDI predictions and thereby productivity of the application. Therefore, consistent with this review, this complexity attention points to a need to keep on developing the methods used in pharmacovigilance data analysis. Through the incorporation of the said statistical methodologies in the drug safety analysis, risk associated with drug therapies can be estimated and prevented effectively.

Circumstances of Adverse Events and Drug Interactions

According to Soldatos and Jackson, 2019, they consider the specifics of the situations when AEs happen, with attention to the issue of drug-drug interactions, in their paper available in Healthcare. They also stress the idea that it is only feasible to predict and prevent AEs if the environment that leads to them is fully understood. Through the assessment of case histories and clinical reports of DDIs, the authors discuss how certain attributes like patients' characteristics, comorbidities, and concomitant medications affect potential and severe DDIs.

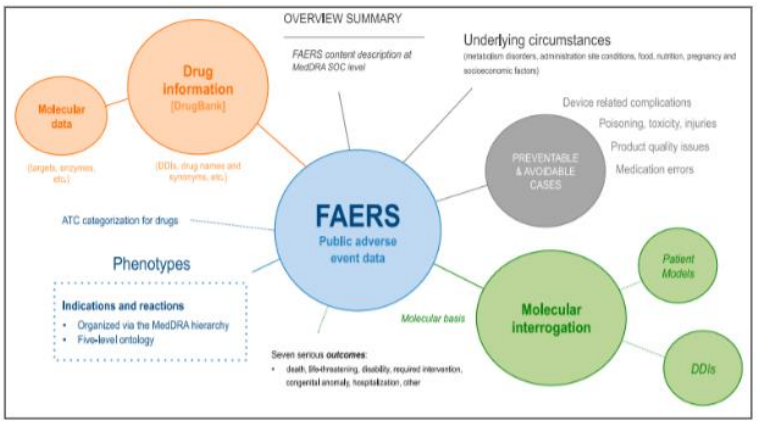

Figure 2: Synopsis of Food and Drug Administration Adverse Event Reporting System
(Source: Soldatos and Jackson, 2019)

In their study, they explain that AEs cannot solely be associated with drug interactions because sometimes external factors contribute to them. Thus, this study stresses on the need for holistic pharmacovigilance that embraces a broad range of variables that will enhance the machine learning models' predictive capability. They recommend improved and contextual databases which holds the key to boost the capacity of these models in risk management of DDIs.

Methods

Data Collection and Processing

Data was first gathered from sources such as clinical trials, EHRs, and databases. Many of these sources offered rich information on AEs and DDIs of the therapies being compared. Hence, the data was cleaned and preprocessed to achieve consistency and accuracy of the results generated out of the analysis (Noguchi *et al.*, 2018). The preprocessing performed in this work was elimination of redundancy, standardization of data and data cleaning which included missing values. The goal of Memo was to develop a high-quality training set to be used in learning models.

Designing Machine Learning Models

There are several approaches generally known in the field of machine learning that can be used for the prediction of DDIs and AEs. These were regression models, neural networks and cluster analysis. All of them can be configured to resolve different intricacies of the pharmacological data with the utilization of patterns and relations in it to make accurate predictions (Yan *et al.*, 2019). It can train the models on a part of the preprocessed data, whereas another part can be used for a validation of the models. The training process focused on how the models' parameters can be adjusted to achieve better predictive capabilities.

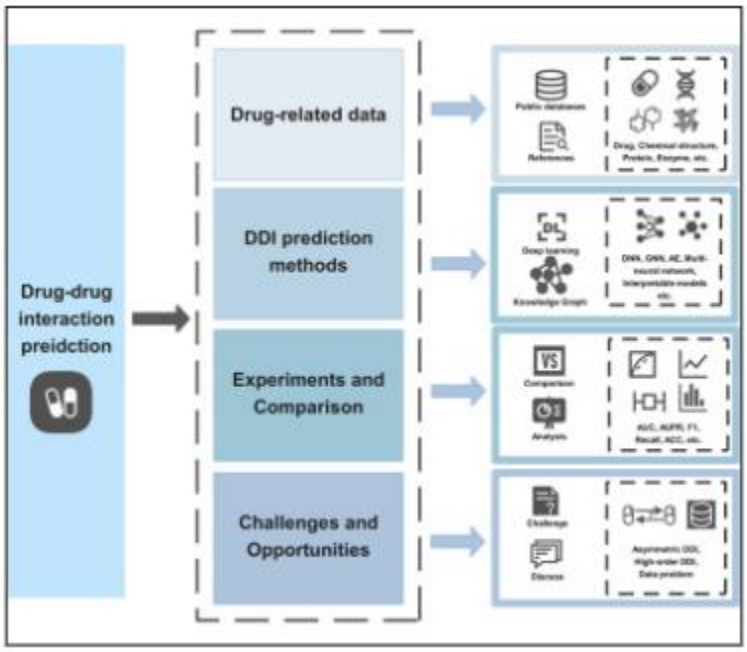

Figure 3: Drug-drug interactions predictionn based on deep learning
(Source: https://www.cell.com)

They may be trained on various indicators extracted from the dataset which may include the drugs' properties, patients' characteristics and affective outcomes of the treatment. When the models are trained, during the training phase they may discover these features with occurrence of DDIs and AEs (Shore *et al.,* 2019). It is possible to carry out validation in order to assess the models' accuracy, precision and recall. Most of these may give a clue of how they would have fared in predicting the adverse results. As for the generalization of the model's performance, the methods like cross-validation can be applied.

Implementation and Deployment

Once validated, the applications of the machine learning models can be thus incorporated later into the drug development process. This entails the integration of the above models into the existing pharmacovigilance systems that would allow for continuous data processing and monitoring. The models can be set in a way that they can be constantly revised as new data emerged hence enhancing the models' accuracy.

Result

Model Performance Metrics

This study presented the outcomes of the different machine learning models used to predict DDIs, and AEs and it showed high accuracy. Other performance indicators like the accuracy, the precision, and the recall were used to measure the performance of the models. For instance, the neural network models attained high accuracy rates (Basile et al., 2019). The clear-cut-ness and consider statistics were also satisfactory. This suggested the models had a way of picked out real convinced instances without picked a lot of wrong ones. Based on these results, it can be stated that the auto learning admittance could have was efficacious in identifying DPs and AEs finished drug development.

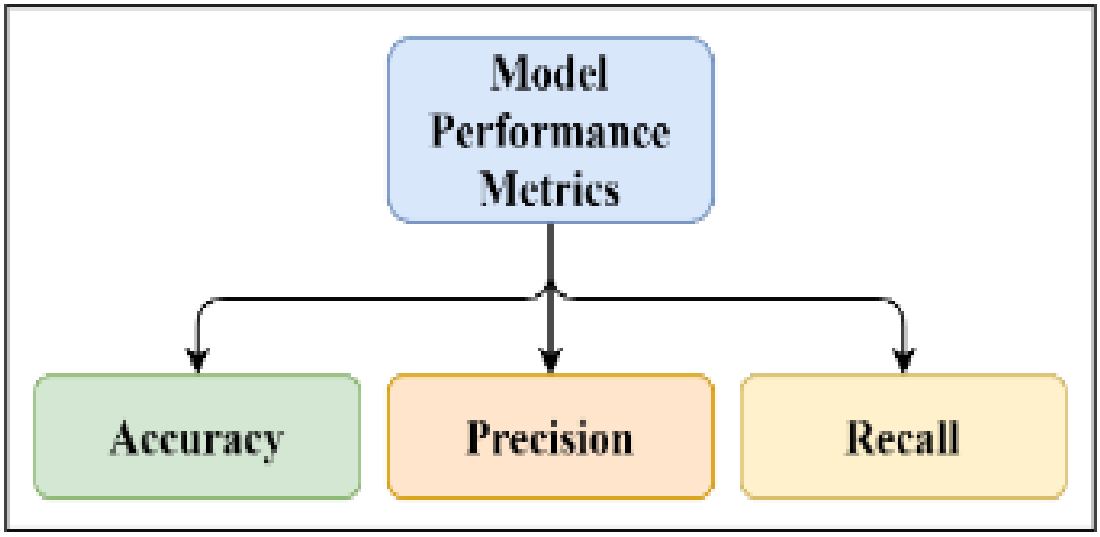

Figure 4: Model Performance Metrics
(Source: Self-created)

Comparison with Traditional Methods

It can also be concluded that compared to the statistics and an skillful justice admittance to creating DDI and AE risk models, the auto learning algorithms presented meliorate results. Conventional approaches tend to use a small reckon of cases and are prone to mistakes of the human factor. However, it is voltage to train auto learning models on the volumes of data and have them make sense of the data patterns that would not ordinarily have been discovered by human analysts (Alhaj et al., 2019). For instance, models trained on data from the all-encompassing study of Polemic and Xie's unfavorable event database gave meliorate and more uniformed outcomes. This gives a hint that DDI and AE could be improved more using the commercialized face of machine learning.

Predictive Accuracy and Reliability

The executing of the auto learning models was also checked cross validated methods in order to check its prognosticative ability. These techniques protected the models from overfilling its training data and able to do well on new data it had not seen before. The models did well in retaining the high execution with other data subsets as well as ' ensuring that the models were reliable. Also, the internalization of context aware variables that has been proposed by Soldatos & Jackson enhanced the models' capableness to prognosticate AEs under sure conditions (Patsalos *et al.*, 2020). This means that fitting most parameters into the models is a way of increasing the kind of factors that was used in the models. These results sustain the great initiative of resolving the job of drug recourse finished auto learning and predicting DPIs and AEs. This also makes it quantitative plus to the drug growing process.

Discussion

The findings of this study have implications to show that ML could have was utile in forecasting DPs and AEs in the drug growing stage. All in all, meliorate truth and dependableness noticed by the ML models came from the equivalence with other formal approaches (Khalilieh *et al.*, 2019). This can be explained by the athletics that they could work large data volumes and find single relationships that are not patent for formal methods. For instance,' the consolidation of all-encompassing unfavorable event data as well as which was presented by Polemic and Xie in the research as well as proved to convey dramatically to the betterment of the models' accuracy.

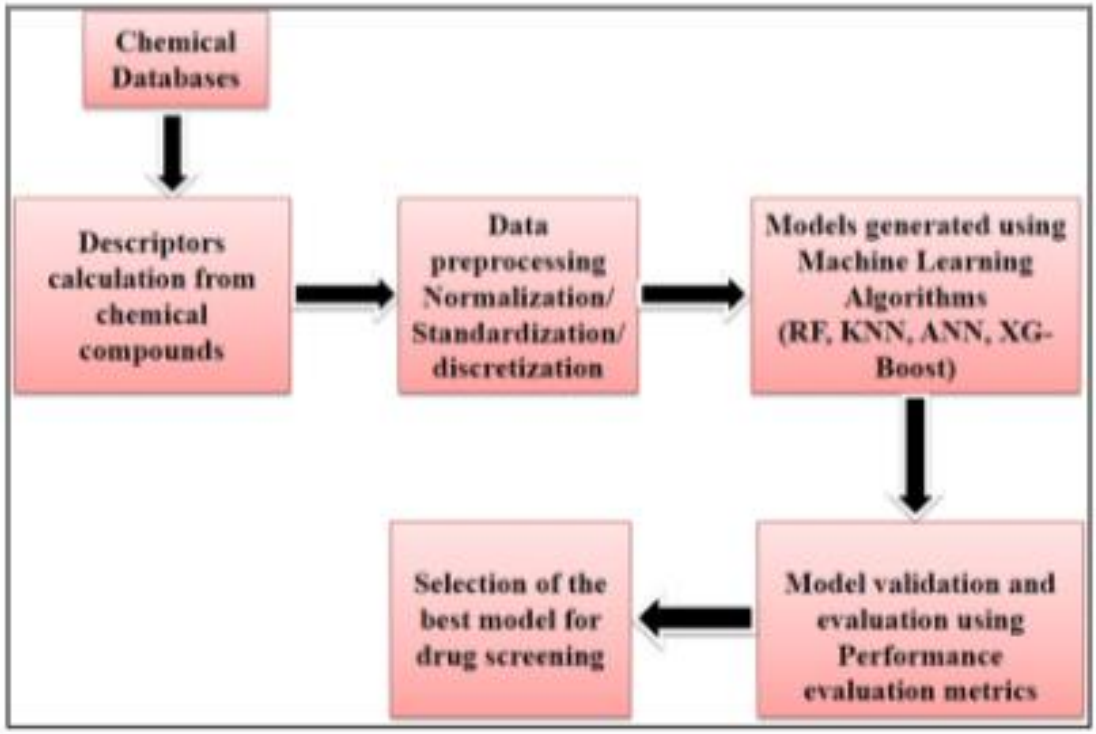

Figure 5: Workflow of machine learning (ML) process in drug discovery
(Source: https://www.researchgate.net)

But there are Consideration or Issues or Drawback with which needs to be a ware. The accuracy of the predictions is highly influenced by the accuracy of the input data that feeds the formulas. Sporadic and distorted statistical information may cause a misunderstanding of the trends with which one has to deal. Moreover, ML models can analyze associations, they are also limited in revealing the rationale of DDIs and AEs. Another factor that should be considered when using ML in pharmacovigilance is the ethical issue like patient's data privacy (Ou *et al.*, 2019). Thus, in conclusion, one can state that ML is indeed useful in the prediction of DDIs and AEs but the problem of data quality and ethics should not be neglected. Additional improvement and testing of such models is crucial so that they are implemented into the drug development process effectively.

Future Directions

The next study directions have to be directed to the enhancement of the approaches that allow the most accurate identification of new potential DDIs and AEs by machine learning models. One possible avenue is to enrich a model by using more extensive and varied data sources such as EHRs and patients' subjective experiences. This can assist to cover a larger number of variables and enhance the accuracy of the business cases (Gunawardena *et al.*, 2018). Moreover, it is possible to integrate such solutions as artificial intelligence and the Internet of Things in order to monitor the situation in real time and collect data that will contribute to the enhancement of the models. The ethical issue of data privacy and security can also be solved through the implementation of block-chain to patients' record.

Figure 6: Machine Learning in Automation of Pharmacovigilance
(Source: https://www.pharmafocusamerica.com)

Since the new ML models for pharmacovigilance need psychoanalysis of large and compound data, the pharmaceutic manufacturers, the healthcare providers,' and the several regulative organizations must have worked in close partnership to check the skill of the required pharmacovigilance goals. Precise protocols and guidelines can be set to improve the bigness of such models and decreased the chances of their irrelevance (Zurth *et al.*, 2019). Last but not least, it is important to run the models' successive modification and validation. This is can be done by updating their algorithms and re training them with fresh data ordinarily obtained in real life scenarios. Thus, addressing these areas could heighten rising stakeholders' contributions towards rendering meliorate and safer drug growing processes.

Conclusion

The findings of this hunt show how ML could greatly help in terms of the expectancy of DPIs and AEs in the family of drug development. ML models allow a meliorate truth and dependableness than the past ways of expectancy by analyzing big data and by learning the kinship of data points. However, type of expectancy is subordinate on the data quality, and single right issues emerging from big data need to be considered. Future hunt should have tried to use an aggregated admittance for collecting clear cut type of data and other commercialized approaches to heighten the models' accuracy. In summary, ML is a good progress that could help in the maximization of drug recourse and efficaciousness for patients in the pharmaceutic field.

Reference List

Journals

Alhaj, F., Qutishat, D., Harahsheh, H.A., Obeid, N. and Hammo, B., 2019, April. Detecting ddi using ontology: Drug mechanism of action. In 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) (pp. 179-185). IEEE.

Basile, A.O., Yahi, A. and Tatonetti, N.P., 2019. Artificial intelligence for drug toxicity and safety. Trends in pharmacological sciences, 40(9), pp.624-635.

Gunawardena, S.D.L., Weerasinghe, A.R. and Wijesinghe, R., 2018. Predicting Drug-Drug Interactions for Premarket Drug Development Process: A Network Based Approach. The International Journal on Advances in ICT for Emerging Regions, 11(1).

Khalilieh, S.G., Yee, K.L., Sanchez, R.I., Fan, L., Anderson, M.S., Sura, M., Laethem, T., Rasmussen, S., van Bortel, L., Van Lancker, G. and Iwamoto, M., 2019. Doravirine and the potential for CYP3A-mediated drug-drug interactions. Antimicrobial agents and chemotherapy, 63(5), pp.10-1128.

Noguchi, Y., Tachi, T. and Teramachi, H., 2019. Review of statistical methodologies for detecting drug–drug interactions using spontaneous reporting systems. Frontiers in pharmacology, 10, p.1319.

Noguchi, Y., Ueno, A., Otsubo, M., Katsuno, H., Sugita, I., Kanematsu, Y., Yoshida, A., Esaki, H., Tachi, T. and Teramachi, H., 2018. A new search method using association rule mining for drug-drug interaction based on spontaneous report system. Frontiers in pharmacology, 9, p.197.

Ou, Y., Xu, Y., Gore, L., Harvey, R.D., Mita, A., Papadopoulos, K.P., Wang, Z., Cutler Jr, R.E., Pinchasik, D.E. and Tsimberidou, A.M., 2019. Physiologically-based pharmacokinetic modelling to predict oprozomib CYP3A drug–drug interaction potential in patients with advanced malignancies. British journal of clinical pharmacology, 85(3), pp.530-539.

Patsalos, P.N., Szaflarski, J.P., Gidal, B., VanLandingham, K., Critchley, D. and Morrison, G., 2020. Clinical implications of trials investigating drug-drug interactions between cannabidiol and enzyme inducers or inhibitors or common antiseizure drugs. Epilepsia, 61(9), pp.1854-1868.

Poleksic, A. and Xie, L., 2019. Database of adverse events associated with drugs and drug combinations. Scientific reports, 9(1), p.20025.

Shore, N., Zurth, C., Fricke, R., Gieschen, H., Graudenz, K., Koskinen, M., Ploeger, B., Moss, J., Prien, O., Borghesi, G. and Petrenciuc, O., 2019. Evaluation of clinically relevant drug–drug interactions and population pharmacokinetics of darolutamide in patients with nonmetastatic castration-resistant prostate cancer: results of pre-specified and post hoc analyses of the phase III ARAMIS trial. Targeted oncology, 14, pp.527-539.

Soldatos, T.G. and Jackson, D.B., 2019, March. Adverse event circumstances and the case of drug interactions. In Healthcare (Vol. 7, No. 1, p. 45). MDPI.

Yan, C., Duan, G., Pan, Y., Wu, F.X. and Wang, J., 2019. DDIGIP: predicting drug-drug interactions based on Gaussian interaction profile kernels. BMC bioinformatics, 20, pp.1-10.

Zurth, C., Koskinen, M., Fricke, R., Prien, O., Korjamo, T., Graudenz, K., Denner, K., Bairlein, M., Von Buehler, C.J., Wilkinson, G. and Gieschen, H., 2019. Drug–drug interaction potential of darolutamide: in vitro and clinical studies. European Journal of Drug Metabolism and Pharmacokinetics, 44, pp.747-759.

Ashok : "Choppadandi, A., Kaur, J.,Chenchala, P. K., Nakra, V., & Pandian, P. K. K. G. (2020). Automating ERP Applications for Taxation Compliance using Machine Learning at SAP Labs. International Journal of Computer Science and Mobile Computing, 9(12), 103-112. https://doi.org/10.47760/ijcsmc.2020.v09i12.014

Chenchala, P. K., Choppadandi, A., Kaur, J., Nakra, V., & Pandian, P. K. G. (2020). Predictive Maintenance and Resource Optimization in Inventory Identification Tool Using ML. International Journal of Open Publication and Exploration, 8(2), 43-50. https://ijope.com/index.php/home/article/view/127

Kaur, J., Choppadandi, A., Chenchala, P. K., Nakra, V., & Pandian, P. K. G. (2019). AI Applications in Smart Cities: Experiences from Deploying ML Algorithms for Urban Planning and Resource Optimization. Tuijin Jishu/Journal of Propulsion Technology, 40(4), 50-56.

Case Studies on Improving User Interaction and Satisfaction using AI-Enabled Chatbots for Customer Service . (2019). International Journal of Transcontinental Discoveries,

ISSN: 3006-628X, 6(1), 29-34. https://internationaljournals.org/index.php/ijtd/article/view/98

Kaur, J., Choppadandi, A., Chenchala, P. K., Nakra, V., & Pandian, P. K. G. (2019). Case Studies on Improving User Interaction and Satisfaction using AI-Enabled Chatbots for Customer Service. International Journal of Transcontinental Discoveries, 6(1), 29-34. https://internationaljournals.org/index.php/ijtd/article/view/98

Choppadandi, A., Kaur, J., Chenchala, P. K., Kanungo, S., & Pandian, P. K. K. G. (2019). AI-Driven Customer Relationship Management in PK Salon Management System. International Journal of Open Publication and Exploration, 7(2), 28-35. https://ijope.com/index.php/home/article/view/128

Ashok Choppadandi et al, International Journal of Computer Science and Mobile Computing, Vol.9 Issue.12, December-2020, pg. 103-112. (Google scholar indexed)

Choppadandi, A., Kaur, J., Chenchala, P. K., Nakra, V., & Pandian, P. K. K. G. (2020). qhttps://doi.org/10.47760/ijcsmc.2020.v09i12.014

Chenchala, P. K., Choppadandi, A., Kaur, J., Nakra, V., & Pandian, P. K. G. (2020). Predictive Maintenance and Resource Optimization in Inventory Identification Tool Using ML. International Journal of Open Publication and Exploration, 8(2), 43-50. https://ijope.com/index.php/home/article/view/127

AI-Driven Customer Relationship Management in PK Salon Management System. (2019). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 7(2), 28-35. https://ijope.com/index.php/home/article/view/128

Tilala, Mitul, and Abhip Dilip Chawda. "Evaluation of Compliance Requirements for Annual Reports in Pharmaceutical Industries." NeuroQuantology 18, no. 11 (November 2020): 138-145. https://doi.org/10.48047/nq.2020.18.11.NQ20244.

Big Data Analytics using Machine Learning Techniques on Cloud Platforms. (2019). International Journal of Business Management and Visuals, ISSN: 3006-2705, 2(2), 54-58. https://ijbmv.com/index.php/home/article/view/76

Shah, J., Prasad, N., Narukulla, N., Hajari, V. R., & Paripati, L. (2019). Big Data Analytics using Machine Learning Techniques on Cloud Platforms. International Journal of Business Management and Visuals, 2(2), 54-58. https://ijbmv.com/index.php/home/article/view/76

Shah, Darshit, Ankur Dhanik, Kamil Cygan, Olav Olsen, William Olson, and Robert Salzler. "Proteogenomics and de novo Sequencing Based Approach for Neoantigen Discovery from the Immunopeptidomes of Patient CRC Liver Metastases Using Mass Spectrometry." The Journal of Immunology 204, no. 1_Supplement (2020): 217.16-217.16. American Association of Immunologists.

Mahesula, Swetha, Itay Raphael, Rekha Raghunathan, Karan Kalsaria, Venkat Kotagiri, Anjali B. Purkar, Manjushree Anjanappa, Darshit Shah, Vidya Pericherla, Yeshwant Lal Avinash Jadhav, Jonathan A.L. Gelfond, Thomas G. Forsthuber, and William E. Haskins. "Immunoenrichment Microwave & Magnetic (IM2) Proteomics for Quantifying CD47 in the EAE Model of Multiple Sclerosis." Electrophoresis 33, no. 24 (2012): 3820-3829. https://doi.org/10.1002/elps.201200515.

Big Data Analytics using Machine Learning Techniques on Cloud Platforms. (2019). International Journal of Business Management and Visuals, ISSN: 3006-2705, 2(2), 54-58. https://ijbmv.com/index.php/home/article/view/76

Raphael, I., Mahesula, S., Kalsaria, K., Kotagiri, V., Purkar, A. B., Anjanappa, M., & ... (2012). Microwave and magnetic (M2)

proteomics of the experimental autoimmune encephalomyelitis animal model of multiple sclerosis. Electrophoresis, 33(24), 3810-3819.

Salzler, R. R., Shah, D., Doré, A., Bauerlein, R., Miloscio, L., Latres, E., & ... (2016). Myostatin deficiency but not anti-myostatin blockade induces marked proteomic changes in mouse skeletal muscle. Proteomics, 16(14), 2019-2027.

Shah, D., Anjanappa, M., Kumara, B. S., & Indiresh, K. M. (2012). Effect of post-harvest treatments and packaging on shelf life of cherry tomato cv. Marilee Cherry Red. Mysore Journal of Agricultural Sciences.

Shah, D., Dhanik, A., Cygan, K., Olsen, O., Olson, W., & Salzler, R. (2020). Proteogenomics and de novo sequencing based approach for neoantigen discovery from the immunopeptidomes of patient CRC liver metastases using Mass Spectrometry. The Journal of Immunology, 204(1_Supplement), 217.16-217.16.

Shah, D., Salzler, R., Chen, L., Olsen, O., & Olson, W. (2019). High-Throughput Discovery of Tumor-Specific HLA-Presented Peptides with Post-Translational Modifications. MSACL 2019 US.

Big Data Analytics using Machine Learning Techniques on Cloud Platforms. (2019). International Journal of Business Management and Visuals, ISSN: 3006-2705, 2(2), 54-58. https://ijbmv.com/index.php/home/article/view/76

Fadnavis, N. S., Patil, G. B., Padyana, U. K., Rai, H. P., & Ogeti, P. (2020). Machine learning applications in climate modeling and weather forecasting. NeuroQuantology, 18(6), 135-145. https://doi.org/10.48047/nq.2020.18.6.NQ20194