Original Article

Machine learning for decision making in medicine and healthcare

Borislava Toleva, Ivan Ivanov
Faculty of Economics and Business Administration, Sofia University "St. Kl. Ohridski"
Email: bvrigazova@gmail.com
DOI: 10.47750/pnr.2023.14.S02.295

Abstract

In this research we summarize how machine learning algorithms can be used for decision making that can affect health policies. We present modified ANOVA algorithm for identifying marker leukemia genes that allows deeper examination of genes subsets that can increase the risk of developing leukemia. The algorithm uses the ANOVA, the bootstrap and classification to provide an insight whether a particular group of genes affects cancer development. So, medical practitioners can select a group of leukemia genes, test it by the algorithm and further decide whether to examine the group in medical test or switch a gene in the subset. We also present algorithms to outline factors that affect covid geographical distribution and use of vaccines. Some of our key findings are that leukemia genes can be ranked by importance and in rare cases mutations in less important genes can also lead to leukemia development. In terms of covid, we find that the economic development of a country can be related to the willingness of people to vaccinate.

INTRODUCTION

Machine learning (ML) algorithms have become an essential part of everyday life. Their ability to process large quantity of data has allowed society to get deeper knowledge of the world. Health and economics are only part of the fields where machine learning algorithms have become irreplaceable for decision making. Machine learning is used in the healthcare sector for developing algorithms for diseases prevention (Spathis et.al., 2022), exploring the consequences from climate change on health (Berrang-Ford et.al., 2021) and early detection of disease (Liu et. al., 2021), (Bouazza et. al., 2018). Machine learning algorithms have also become an integral part of statistics handbooks for economics (Duarte et.al.,2020) and medicine (Chang, 2020).

A very important topic in medicine is detection of marker genes for cancer detection, prevention and treatment. Genes have been collected on microchips, so they can be classified as microarray data (Hambali et. al., 2021). Most of the algorithms perform genes selection by feature selection, feature extraction or filters (Hambali et. al., 2021). Cluster analysis (Svrakic et. al., 2003), machine learning (Rendón et. al., 2020), (Basgall et. al., 2019), deep learning (Reyes-Nava et. al., 2018) and classification (Cao, 2015) have been the most common approaches used to detect cancerous genes. However, there is a gap in all cancer research. Machine learning models identify diverse sets of marker genes. Some genes are identified by various machine learning models, while others are identified as marker in one research but not in the other. Many examples of this issue can be found in the above papers. To help medical researchers decide which genes to test as marker in a laboratory, we propose the ANOVA-Bootstrap-Classification model (Vrigazova & Ivanov, 2020). This would be one of the topics in the chapter.

Another important topic where ML algorithms can be used is modelling economic effects from healthcare sector changes. This topic has become extremely popular since 2020, when covid19 became a constant part of our lives. The virus led to policymaking that needed to handle the trade-off between health and poverty (Stiglitz, 2021). Protecting people's health came at the cost of bankrupts of small and medium-sized businesses (Hertz-Palmor et. al., 2021), job loss (Vieira et. al., 2021) and stress on the financial and health care system (Halley et. al., 2021), (Firano et. al., 2022). Therefore, predicting the spread and severity of covid19 and how to handle its effects have become another topic modelled by ML (Shahin et.al., 2022). Telemedicine is another field where ML algorithms have become an important tool to detect and treat covid (Schünke et. al., 2022). Therefore, in this chapter we also aim to show other applications of ML to model covid19 and its effects on the economy.

Next section clarifies the machine learning algorithms we use to achieve the two goals of this chapter. Section Results focuses on our findings and comparison with other authors. The last section concludes.

Methodology:

The analysis of variance (ANOVA) is a powerful statistical tool (James et. al., 2021). In this chapter we use ANOVA to analyze the relationship between the stage of economic development of the country and covid vaccinates rates. ANOVA is used to test for statistical significance in the mean value of indicators grouped (Y variable) by a particular criterion (X variable) (Tuckey, 1949). Thus, providing an answer to the question 'How does the grouping criterion affects the indicator'. We use a public dataset (https://data.covid19taskforce.com/data/tables), where we define the X variable as Income Group and the Y variable represent one of various quantitative covid vaccination indicators. We examine four groups of economic development – countries with low income, countries with lower-middle income, countries with upper-middle income and countries with high income. We run one-way ANOVA, where all income groups are X, while the particular covid vaccination indicator is Y.

Machine learning models combined with statistical models like the ANOVA can become a powerful tool for data analysis. For instance, when ANOVA is combined with classification models and the bootstrap (Vrigazova, Ivanov, 2020), (Vrigazova, 2020) to rank important genes, whose mutations are likely to cause leukemia. Unlike previous leukemia research (Cao, 2015), (Gao, 2017), we introduce the idea that leukemia genes should be ranked by importance to select combination of genes for further laboratory experiments. This would make medical research much more efficient.

The ANOVA-Bootstrap- Classification (Vrigazova, Ivanov, 2020), (Vrigazova, 2020) can be summarized in several steps (figure 1). The algorithm ranks all genes based on their variance (ANOVA). This ranking is constant at all stages of the algorithm. After all steps are completed, the output contains the accuracy for a number of genes starting from the first ANOVA ranked gene and ending with the least important. For example, the first group contains the first 200 genes ranked by the ANOVA model, second – the first 500, etc. and the accuracy achieved by the classification model for the particular group of genes. The algorithm displays all groups of genes and their accuracy from the respective classification model on a chart. The group of genes that has the highest accuracy is selected. For example, if the number 250 is chosen, this means that the top 250 genes with the highest ANOVA scores are selected in the classification model. The procedure on figure 1 can be repeated many times until a smaller group of genes is selected depending on the purpose of the research.

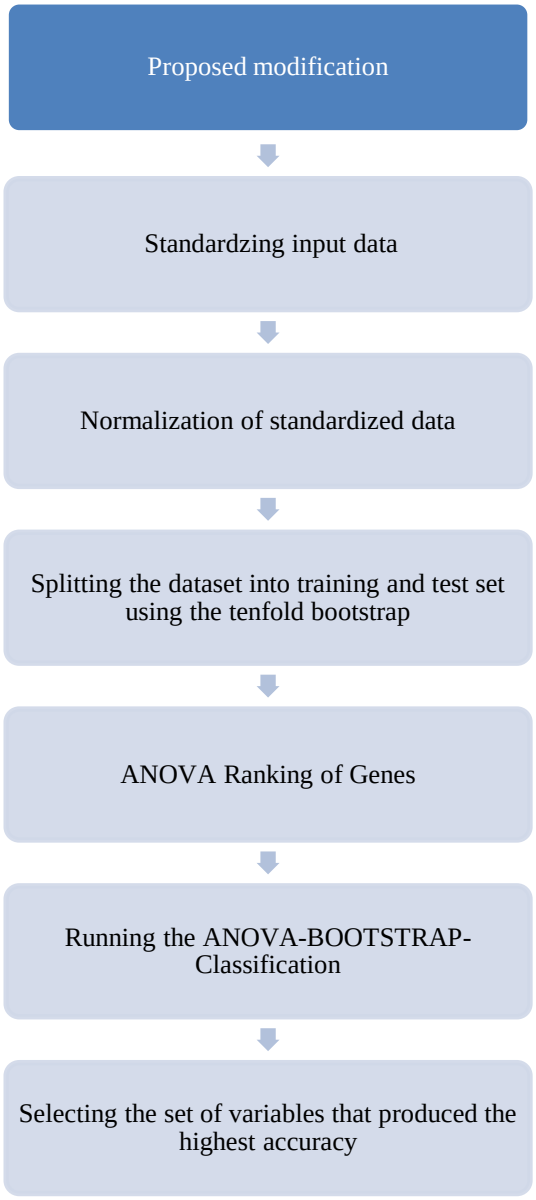


Figure 1: The ANOVA-Bootstrap-Classification. Source: the authors

The advantage of this novel algorithm is that it can be used for both identifying marker genes /genes, whose mutations can potentially lead to cancer/ and applying genes scenarios for laboratory research. The bootstrap procedure improves the model performance (Vrigazova, Ivanov, 2020), while the ANOVA allows grouping of genes to fulfil the purpose of the research. Classification models include the logistic regression (LR), Naïve Bayes (NB), support vector machines (SVM), linear discriminant analysis (LDA) and the k-nearest neighbors (KNN) and decision trees (DT).

Next section discusses our results.

Results and Discussion

The leukemia dataset contains 7130 genes. The ANOVA ranks them by importance, starting with the genes whose mutations are more likely to cause leukemia. Table 1 presents the first 100 most important genes and their ANOVA scores, defined by our model. Rank 1 means the most important genes, rank 100 – less important gene than the previous ranks.

Table 1: ANOVA scores of the First 100 Most Important Leukemia Genes

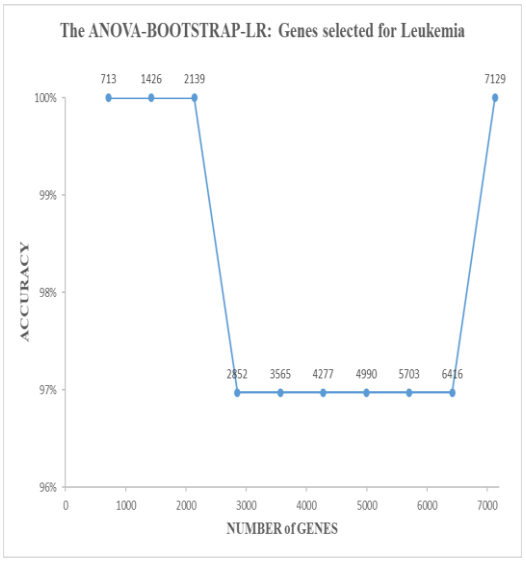
Rank	Gene	F-score	Rank	Gene	F-score	Rank	Gene	F-score
1	M84526	79.6	**34**	M28170	31.9	**67**	L19437	24.3
2	M22960	79.3	**35**	X87613	31.3	**68**	Y00787	24.2
3	X95735	77.4	**36**	Y00282	29.6	**69**	M34423	23.9
4	U46499	77.1	**37**	M38449	29.2	**70**	U19878	23.9
5	M23197	75.6	**38**	D88422	29.1	**71**	U50327	23.6
6	M27891	67.7	**39**	M98045	28.9	**72**	AB002559	23.3
7	M93056	62.4	**40**	J00148_cds2_f	28.3	**73**	Y00433	23.2
8	M63138	59.9	**41**	M32304	28.3	**74**	J05243	22.9
9	L09209	58.4	**42**	U19713	28.3	**75**	X13839	22.6
10	L09717	51.5	**43**	X52056	27.8	**76**	U07139	22.6
11	M14636	46.8	**44**	HG2562-HT2658	27.7	**77**	M11147	22.6
12	M63959	45.7	**45**	X59417	27.6	**78**	X62320	22.5
13	M19507	42.7	**46**	M21535	27.2	**79**	M14328	22.3
14	Y07604	42.1	**47**	J02783	27.1	**80**	U50136	22.3
15	U10868	42.0	**48**	M59820	27.1	**81**	X78669	22.1
16	Z29067	41.4	**49**	U30255	27.1	**82**	X04828	22.0
17	X64072	41.0	**50**	M80647	27.0	**83**	X97267	21.9
18	X61587	40.8	**51**	U22376_cds2	26.8	**84**	M11722	21.8
19	L41559	39.4	**52**	D29643	26.7	**85**	X91911	21.1
20	M92287	39.3	**53**	L11669	26.6	**86**	X51521	21.1
21	D49950	38.6	**54**	M13452	26.4	**87**	X80230	21.0
22	X16546	38.3	**55**	X98411	26.2	**88**	M65214	20.9
23	M62762	38.1	**56**	S82470	26.1	**89**	J03801_f	20.9
24	U05259	37.6	**57**	X59711	25.8	**90**	K01396	20.7
25	U05572	37.6	**58**	HG1612-HT1612	25.7	**91**	X05908	20.7
26	X62654	36.9	**59**	L21954	25.7	**92**	S50223	20.6
27	M16038	35.8	**60**	AF005043	25.6	**93**	X14008_f	20.4
28	M15395	34.0	**61**	X17042	25.5	**94**	M19508_xpt3	20.2
29	D14664	33.9	**62**	M83652	24.9	**95**	M31211	20.2
30	M31523	33.6	**63**	X66401_cds1	24.8	**96**	U14588	19.9
31	M55150	32.4	**64**	D31886	24.7	**97**	U53468	19.8
32	M96326	32.1	**65**	M84371	24.6	**98**	L06797	19.8
33	X12447	31.9	**66**	M89957	24.5	**99**	X84194	19.7
						100	X15414	19.7

Source: authors' research

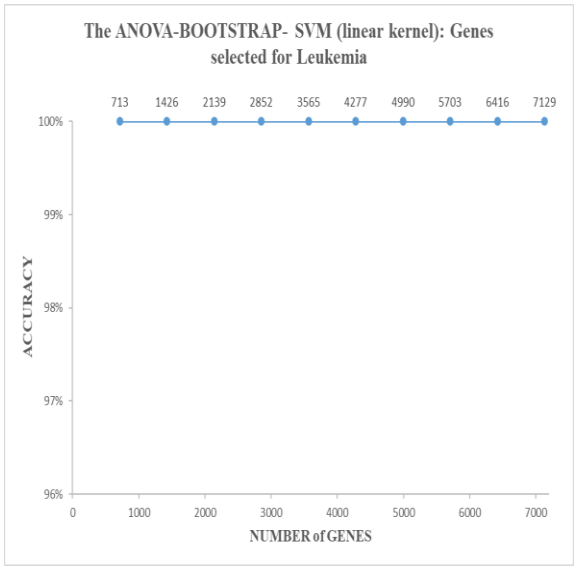
However, using all genes is not efficient in medical research. The aim is to select a smaller group of genes that would be representative for developing fast leukemia blood tests. Feature selection is the next step in our model. Usually, blood tests are based on a few genes, for example 5 or 10. ML models result in selecting a variety of genes. Some authors propose 4,13,80 genes (Bouazza et.al., 2018), others – about 100 genes (Huerta, 2008). Based on the ML algorithm the genes selected vary from

research to research. Although some genes are selected by many algorithms, the genes that are different from each research can bring doubt whether we the selected genes are valid.

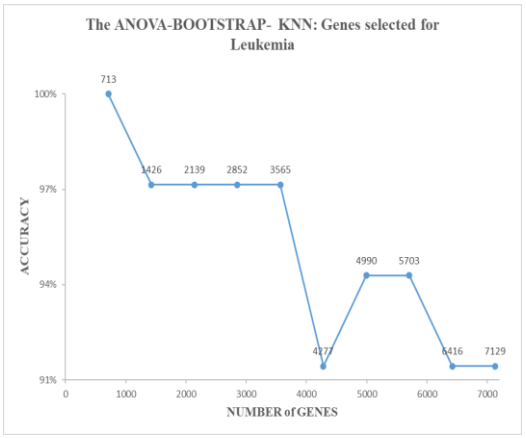
The advantage of our algorithm is that the ANOVA genes ranking in table 1 remains constant despite the classification model. The next step is to select the combination of the ANOVA ranked genes that would result in the best accuracy and the smallest number of genes. Figure 1 presents graphically the output of each ANOVA-Bootstrap classification model. Based on each chart, we select the narrowest combination of genes that result in the highest accuracy for each model.



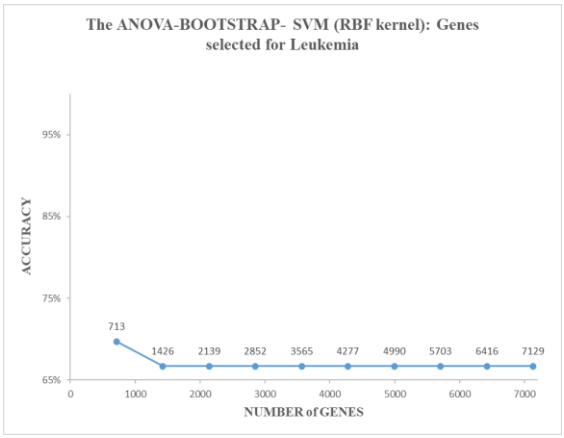
a



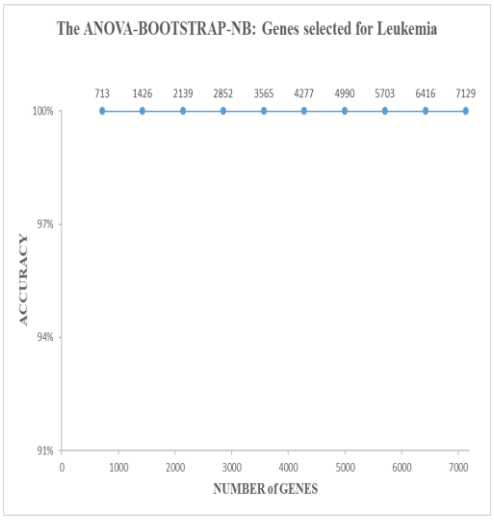
b



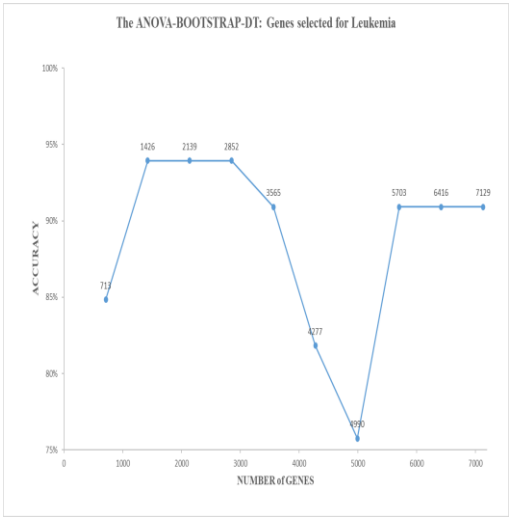
c



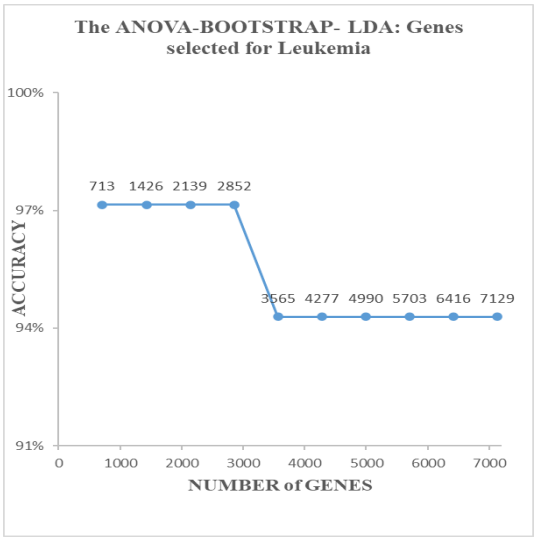
d



e



f



g

Figure 1: ANOVA-Bootstrap-Classification

The ANOVA-Bootstrapped LDA, DT and SVM (with RBF kernel) do not result in 100% accuracy like the rest of the models. We can exclude them from our analysis, as other models produce higher accuracy than those. Figure 1 a,b,c and e show the smallest number of genes that each model requires to achieve 100% accuracy. In the case of the ANOVA-Bootstrap-LR (a) the first 713 most genes are selected. This means that the model requires the first 100 genes showed in table 1 and the remaining 613 ANOVA ranked genes. It is important to note that the ANOVA ranking changes the initial place of the genes in the dataset. The ANOVA-Bootstrap-SVM(linear kernel) (b) also uses the same 713 genes as the LR model to achieve 100% accuracy. Similar is the situation in the KNN model (c) and the Naïve Bayes in point c.

Therefore, our experiments confirmed that using the first 713 most important genes as ranked by the ANOVA can achieve 100% accuracy in predicting leukemia cases. Unlike most academic research, the output of our algorithm results in the same genes regardless of the classification model used. As a result, medical researchers can have a preselected subset of leukemia genes that can first be examined in laboratory. Thus, time for research can be saved as laboratory researchers have a predefined idea of the relationship between genes' mutations and leukemia. The subset of important genes can be updated each time more data are included in the model by rerunning the algorithm. So, we offer solution to the controversial ML output presented in academic research.

To make medical research more efficient, we rerun the ANOVA-Bootstrap-classification on the subset of 713 most important genes. We aim to select a smaller subset of genes that can be used for initial research. The ANOVA-Bootstrapped LR and SVM (linear kernel) resulted in 100% accuracy using only 10 genes. The ten most important genes selected by our algorithm are genes ranked 1-10 in table 1, namely M84526, M22960, X95735, U46499, M23197, M27891, M93056, M63138, L09209 and L09717. Those genes are also selected as marker by other authors (Bouazza, 2018), (Eising, 2013), (Huerta, 2018), although our research is the first ranking them.

Our results suggest that medical researchers should start examining the relationship between the ten most important genes and leukemia in laboratory conditions. Another advantage of our algorithm is that medical researchers can use it to form a control subset, where to replace a particular gene and examine its effects on leukemia development. For instance, take the subset of 10 most important genes and replace the tenth genes by the least important gene. Then examine whether the subset of genes together change their relationship with leukemia development, what the influence of every gene is. As a result, they can select genes for medication development and screening tests in a much more efficient way. A faster breakthrough can then happen in leukemia prevention and treatment. Also, the proposed algorithm can be applied to other types of cancer.

Another important application of the ANOVA in healthcare is modelling the relationship between covid and economic development. We examine the relationship between the stage of economic development of the country and the vaccination rates. Our findings suggest that the number of doses delivered, the current number of cases and the average daily vaccination rate observed in the past few days do not affect vaccination rates regardless of the country's stage of economic development. These results are important for economic decision making in light of covid19. On the hand, they show that people's decision to vaccinate is not affected by the covid19 cases and vaccination statistics in both rich and poor countries. Therefore, economic development may affect the decision to vaccinate against covid19 in other ways. For instance, the percentage of population fully vaccinated, percentage of population vaccinated at least one dose, maximum daily vaccination rate observed so far and the administered doses as percentage of delivered exhibit statistically significant difference in the means when grouped by income.

Administered doses as percentage of delivered defines how many vaccines have been used for vaccination as percentage of the total number delivered. It can be used to monitor vaccines utilization rates. Our study suggests that low-income countries tend to utilize a bigger percentage of vaccines than upper-middle income countries and rich countries. Low-to-middle income countries also have higher administered doses as percentage of delivered than high-income countries. This trend is not visible when comparing countries with similar economic development, e.g.low-income and low-upper middle-income countries. The bigger the difference between the stage of economic development, the more likely it is high vaccination rates to be present in poor and low-upper middle-income countries than in rich countries. A higher percentage of doses has been utilized in poor countries than in rich countries. So, more vaccines can be redirected to poor countries than to rich and achieve a more efficient distribution of vaccines. Therefore, the issue with many unused vaccines in richer countries may be solved by ordering a smaller number of vaccines.

The maximum daily vaccination rate observed so far (per 100 people) is another statistically significant variable in our model. Low-income countries exhibit a higher mean difference in the maximum daily vaccination rate observed so far than high-income countries and upper-middle income countries. Citizens of poorer countries seem to be more willing to get vaccinated than those in upper-middle income countries. The richer the country becomes, the less observable is the difference between citizens' attitude towards covid vaccination. This finding also has decision making implications when planning the delivery and distribution of vaccines. Covid measures can also be affected by the citizens' attitudes towards covid vaccination.

High-income countries tend to have by about 60.5% lower percentage of population vaccinated at least one dose compared to low-income countries. They also exhibit by about 31.1% lower percentage than low-middle income countries and by about 22.9% lower than upper-middle income countries. Citizens of richer countries are more likely to wait before getting a first covid dose. One possible reason is that their wealth and health may be less dependent on vaccines as they may have easier access to treatment and their jobs may be at a lower risk of lost. Covid vaccines seem to be both an economic and health measure in poor countries, while it may be more of a health than an economic measure in richer countries. So, covid vaccination can be a tool to fight poverty in poorer countries.

Similar findings can be found in the trend of getting a full cycle of covid vaccination. The richer the country, the lower proportion of fully vaccinated citizens has been observed. However, this trend is observed at all stages of economic development. However, the percentage of people with at least one dose tend to be similar in countries with similar economic development. This is not the case in terms of fully vaccinated, where the difference in means is statistically significant at all stages of economic development. It seems that people from middle-income countries value a full vaccination cycle more than

getting only one dose. Again, policymakers can consider these results when deciding the rules and time periods for further vaccination.

As shown in the examples above, statistical models and machine learning can be important tools for policymaking and economic decisions. On the hand, they enable us to model given conditions and observe what factors can be considered important for the scenario. In the case of cancer, researchers try to identify a small set of genes, whose mutations would lead to cancel development. Machine learning models are an effective tool to do that. But such outcome would have a far greater impact on the social system than previously thought. Genes selection can lead to more effective and cheaper cancer treatments. Also, a faster diagnose and improved quality of life while being treated. Such an outcome would improve the patient's ability to work, contribute and take advantage of the economic system. Cheaper treatment would allow poorer people to have access to better cancer treatment, thus changing income and social inequality, etc. By modelling medical issues, machine learning algorithms can have a much broader impact on people's health, economics and the social system. Therefore, machine learning for decision makers in healthcare plays a vital role.

ML models can also assist policymakers in structuring economic, health and social rules in times of crisis like the case with covid19. Not only do they help in predicting the pattern of the disease and the restrictions to tackle it, but also in helping policymakers decide on effective strategies to convince people to vaccinate. Rules and covid statistics are no longer enough to encourage people to vaccinate. ML algorithms can provide a deeper understanding of the psychological and economic incentives that people have when deciding on covid vaccination. Also, as we showed machine learning can lead to optimal distribution of covid vaccinations and decrease the number of unused vaccines that should be donated.

Conclusion

In this chapter we summarized some of our recent experiments in the field of healthcare and machine learning to give additional examples of ML models in policymaking in the healthcare. More examples would continue to evolve as ML models constantly improve and continue have a vital role in healthcare. The reason for this is that these models have been the most adaptive and flexible to model complex systems with changing assumptions. As a result, machine learning has now become one of the main tools for healthcare policymakers, thus making the role of the researcher more important than before.

REFERENCES

1. Spathis D., Perez-Pozuelo I., Marques-Fernandez L., Mascolo C., Breaking away from labels: The promise of self-supervised machine learning in intelligent health, Patterns, Volume 3, Issue 2, 2022, ISSN 2666-3899, https://doi.org/10.1016/j.patter.2021.100410.
2. Berrang-Ford L., Sietsma A., Callaghan M., Minx J., Scheelbeek P., Haddaway H., Haines A., Dangour A., Systematic mapping of global research on climate and health: a machine learning review, The Lancet Planetary Health, Volume 5, Issue 8, 2021, Pages e514-e525, SSN 2542-5196, https://doi.org/10.1016/S2542-5196(21)00179-0.
3. Liu Y., Chokka S., Cao B., Chokka P., Screening for bipolar disorder in a tertiary mental health centre using EarlyDetect: A machine learning-based pilot study, Journal of Affective Disorders Reports, Volume 6, 2021, ISSN 2666-9153, https://doi.org/10.1016/j.jadr.2021.100215.
4. Bouazza S., Auhmani K., Zeroual A.,Hamdi N., Selecting significant marker genes from microarray data by filter approach for cancer diagnosis, Procedia Computer Science, Volume 127, 2018, Pages 300-309
5. Chang A., Chapter 5 - Machine and Deep Learning, Editor(s): Anthony C. Chang, Intelligence-Based Medicine, Academic Press, 2020, Pages 67-140, ISBN 9780128233375, https://doi.org/10.1016/B978-0-12-823337-5.00005-6.
6. Duarte V., Duarte D., Fonseca J., Montecinos A., Benchmarking machine-learning software and hardware for quantitative economics, Journal of Economic Dynamics and Control, Volume 111, 2020, 103796, ISSN 0165-1889, https://doi.org/10.1016/j.jedc.2019.103796.
7. Lee J., Choi I., Jun C., An efficient multivariate feature ranking method for gene selection in high-dimensional microarray data, Expert systems with application, 2021, 166:11, p. 3971
8. Hambali M., Oladele Tinuke, Adewole K., Microarray cancer feature selection: Review, challenges and research directions, International Journal of Cognitive Computing in Engineering, Volume 1, 2020, Pages 78-97
9. Basgall, M. J., Hasperué, W., Naiouf, M., Fernández, A., & Herrera, F. (2019). An analysis of local and global solutions to address big data imbalanced classification: A case study with SMOTE preprocessing. In Proceedings of the conference on cloud computing and big data (pp. 75–85).
10. Svrakic, N. M., Nesic, O., Dasu, M. R. K., Herndon, D., & Perez-Polo, J. R. Statistical approach to DNA chip analysis. Recent Progress in Hormone Research, 2003, 58, 75–94.
11. Rendón, E., Alejo, R., Castorena, C., Isidro-ortega, F. J., & Granda-gutiérrez, E. E. Applied sciences data sampling methods to deal with the big data multi-class imbalance problem. Applied Science, 2020, 10(1276), 1–15. 10.3390/app10041276.
12. Reyes-Nava, A., Sánchez, J. S., Alejo, R., Flores-Fuentes, A. A., & Rendón-Lara, E. (2018). Performance analysis of deep neural networks for classification of gene-expression microarrays. In Proceedings of the Mexican conference on pattern recognition (pp. 105 115).
13. Cao J., Zhang L., Wang B.,Li F., Yang J., A fast gene selection method for multi-cancer classification using multiple support vector data description, Journal of Biomedical Informatics, Volume 53, 2015, Pages 381-389

14. Halley M., Stanley T., Maturi J., Goldenberg A., Bernstein J., Wheeler M., Tabor H., “It seems like COVID-19 now is the only disease present on Earth”: living with a rare or undiagnosed disease during the COVID-19 pandemic, Genetics in Medicine, Volume 23, Issue 5, 2021, Pages 837-844, https://doi.org/10.1038/s41436-020-01069-7
15. Hertz-Palmor N., Moore T., Gothelf D., DiDomenico G., Dekel i., Greenberg D., Brown L., Matalon N., Visoki E., White L., Himes M., Schwartz-Lifshitz M., Gross G., Gur R., Gur R., Pessach I., Barzilay R., Association among income loss, financial strain and depressive symptoms during COVID-19: Evidence from two longitudinal studies, Journal of Affective Disorders, Volume 291, 2021, Pages 1-8, https://doi.org/10.1016/j.jad.2021.04.054.
16. Firano A., Adib F., COVID-19 and cross-border contagion: Trade and financial flows, Research in Globalization, Volume 4, 2022, https://doi.org/10.1016/j.resglo.2022.100082.
17. Vieira K., Potrich A., Bressan A., Klein L., Loss of financial well-being in the COVID-19 pandemic: Does job stability make a difference?, Journal of Behavioral and Experimental Finance, Volume 31, 2021, https://doi.org/10.1016/j.jbef.2021.100554 .
18. Stiglitz J., The proper role of government in the market economy: The case of the post-COVID recovery, Journal of Government and Economics, Volume 1, 2021, https://doi.org/10.1016/j.jge.2021.100004.
19. Shahin O., Alshammari H., Taloba A., El-Aziz R., Machine Learning Approach for Autonomous Detection and Classification of COVID-19 Virus, Computers and Electrical Engineering, Volume 101, 2022, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2022.108055.
20. Schünke L., Mello B., Costa C., Antunes R.,Rigo S., Ramos G., Righi R., Scherer J., Donida B., A rapid review of machine learning approaches for telemedicine in the scope of COVID-19, Artificial Intelligence in Medicine, Volume 129, 2022, ISSN 0933-3657, https://doi.org/10.1016/j.artmed.2022.102312.
21. James G., Witten D., Hastie T., Tibshirani R., An Introduction to Statistical Learning with Applications in R, Springer, 2021, https://www.statlearning.com/
22. Tukey, John (1949). "Comparing Individual Means in the Analysis of Variance". Biometrics. 5 (2): 99–114.
23. Multilateral leaders task force on covid19, accessed May 2022, https://data.covid19taskforce.com/data/tables
24. Vrigazova, B.P., Detection of Malignant and Benign Breast Cancer Using the ANOVA-BOOTSTRAP-SVM, Journal of Data and Information Science, 2020, 5(2), pp. 62–75
25. Vrigazova, B.P., Ivanov, I.G., The bootstrap procedure in classification problems, International Journal of Data Mining, Modelling and Management, 2020, 12(4), pp. 428–446
26. Gao L.,Ye M., Lu X., Huang D., Hybrid Method Based on Information Gain and Support Vector Machine for Gene Selection in Cancer Classification, Genomics, Proteomics & Bioinformatics, Volume 15, Issue 6, 2017, Pages 389-395
27. Huerta E., Duval E., Hao J., Fuzzy Logic for Elimination of Redundant Information of Microarray Data, Genomics, Proteomics & Bioinformatics, 2008, Volume 6, Issue 2, Pages 61-73
28. Eisinga R., Breitling R.,Heskes T., The exact probability distribution of the rank product statistics for replicated experiments, FEBS Letters, 2013, Volume 587, Issue 6, Pages 677-682