Cinsiyet Ücret Açığı ve Etkilerinin Derin Öğrenme, Makine Öğrenimi ve Veri Madenciliği İle Analizi

A Deep Learning, Machine Learning and Data Mining Analysis of Gender Pay Gap and Its Influences

Özerk YAVUZ
Dr. Öğretim Üyesi, Nişantaşı Üniversitesi
ozerk@alumni.bilkent.edu.tr
dr.ozerk@gmail.com
https://orcid.org/0000-0002-1371-688X

ÖZET

Tarih boyunca çeşitli cinsiyetler, ırklar, etnik gruplar, çalışanlar, çeşitli cinsel tercih ve yönelimlere sahip bireyler, farklı renklere, yaşam tarzlarına ve tercihlere sahip bireyler, konuşma ve düşünce özgürlüğü, geniş bir yelpazedeki birçok insan ve hayvan toplumun üyeleri ve yasal çerçeve ve anayasalarda haklarını kanunla kazandılar. İnsan hakları, hayvan hakları, işçi hakları, kadın hakları, kadın erkek eşitliği, seçme ve seçilme hakkı, ifade ve düşünce özgürlüğü, insan ırkının zaman içinde yaşadığı ve gördüğü bu hak ve gelişmelerden bazılarıdır. Ancak geçmişte olduğu gibi bu coğrafyaların ve iş ortamlarının bazılarında hala bazı ayrımcılık, damgalama ve benzer sorunlar yaşanması talihsizliktir. Toplumun farklı paradigmaları, değişimler ve yeni gelişmeler hayatımıza girerken, hayatımızda kurallar, düzenlemeler ve yasal çerçevelerle güçlenen birçok değişiklik ve gelişme görüyoruz. Ancak ne yazık ki günümüz dünyasında kaynaklara eşit erişim eksikliği, çifte standardizasyon, fırsat eşitsizlikleri ve çeşitli eşitsizlikler halen mevcuttur. Bu araştırmada, günümüz dünyasının önemli sorunlarından biri olarak sıklıkla dile getirilen cinsiyete dayalı ücret farkı, derin öğrenme, makine öğrenmesi ve veri madenciliği metodolojileri kullanılarak olası etkileriyle birlikte incelenmektedir.

ABSTRACT

Throughout the history several genders, races, ethnic groups, workers, individuals with several sexual preferences and orientations, individuals with different colours, life styles and preferences, freedom of speech and thought, in a broader range many human beings, animals qualified accaptence by different members of the society and won their rights in the legal framework and constitutions by law. Human rights, animal rights, worker rights, women rights, equality of women and men, right to elect and to be elected, freedom of speech and thought are some of these rights and developments human race experienced and saw during time as a result. However as in the past it is misfortune in some of these geographies and business settings still some forms of discrimination, stigmatization and problems alike occur. As the paradigms of the society, change and new developments enter into our lives we see many changes and developments in our lives which are strengthened with rules, regulations and legislative frameworks. However, unfortunately there is still lack of equal access to resources, double standardizations, inequalities of opportunities with several inequalities in today's world. In this research gender pay gap, which is frequently cited as one of the important problems in today's world is examined with potential influences associated with it using deep learning, machine learning and data mining methodologies.

INTRODUCTION

"We should provide equal opportunities for all genders and all human beings regardless of their differences, preferences and life styles for a better society, a sustainable growth and future. For all of us…"

-Asst. Prof. Dr. Özerk Yavuz

"Equality is the soul of liberty; there is, in fact, no liberty without it."

- Frances Wright

"These men ask for just the same thing: fairness, and fairness only. This is, so far as in my power, they, and all others, shall have."

- Abraham Lincoln.

"Equal pay isn't just a women's issue; when women get equal pay, their family incomes rise and the whole family benefits."

-Mike Honda.

"I am fond of pigs. Dogs look up to us. Cats look down on us. Pigs treat us as equals."

- Winston Churchill.

"Race, gender, religion, sexuality, we are all people and that's it."

- Connor Franta.

Throughout the history several genders, races, ethnic groups, workers, individuals with several sexual preferences and orientations, individuals with different colours, life styles and preferences, freedom of speech and thought, in a broader range many human beings, animals qualified accaptence by different members of the society and won their rights in the legal framework and constitutions by law. Human rights, animal rights, worker rights, women rights, equality of women and men, right to elect and to be elected, freedom of speech and thought are some of these rights and developments human race experienced and saw during time as a result. However as in the past it is misfortune in some of these geographies and business settings still some forms of discrimination, stigmatization and problems alike occur. As the paradigms of the society, change and new developments enter into our lives we see many changes and developments in our lives which are strengthened with rules, regulations and legislative frameworks. However, unfortunately there is still lack of equal access to resources, double standardizations, inequalities of opportunities with several inequalities in today's world. In this research gender pay gap, which is frequently cited as one of the important problems in today's world is examined with potential influences associated with it using deep learning, machine learning and data mining methodologies (Kotler, 1991; Kotler, 2000, Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

Today inequalities in business world, inequalities in access to resources, inequalities in pays, gender, color and nationality based discriminations and double standards remain to be some of the problems that is seen in some

countries, geographies and business settings. As stated in literature, gender based income inequalities in business landscape remains to be an important problem in several business settings, countries and geographies which leads to several negative consequences in 21st century. As in other lower income segments, lower income female groups are facing difficulties in affording their living expenses especially in key areas as, education, health, nutrition, housing, self development. They face difficulties in sustaining their lives and the generations that they raise arising from financial difficulties compared to higher income groups of the society. Since they do not afford the education for theirselves and their children, these generations can not have the necessary education to develop capacities and capabilities to be successful and take place in the job market. Since they are not hired with high incomes in many cases they and their offsprings found theirselves in a deadlock. However when male and female groups are given the same opportunities and income levels, it is seen that both gender types can afford their needs, wants and demands in a more equalistic fashion (Kotler, 1991; Kotler, 2000, Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

In order to prevent pay gaps between genders, discriminations associated with gender, paying for the same job titles according to gender, should be prevented. Double standardizations according to gender and other forms of differences of the individuals should not be applied as part of the societal and organizational culture. Equal opportunities and resources should be dedicated to all genders which would enable the society to be developed more healthy and sustainable while preserving the equality between different gender types and providing equal opportunities to all gender types and differences. All life styles, preferences, paradigms, several gifts given by the mother nature as different physical characteristics or gender should be respected and be secured with rules, regulations and legislative actions for a more sustainable, peacefull society where each member of the society can be granted equal rights, opportunities and resources. Paying and discriminating according to gender, nationality or several differences that human beings have should be one of the bad experiences and misfortunes of the past on dusty shelves, not a contemporary world phenomena (Kotler, 1991; Kotler, 2000, Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

In this research deep learning, machine learning and data mining research methodology composed of supervised and unsupervised methodologies involving deep learning techniques have been applied. Data mining is defined as a methodological approach in quantitative data analysis as indicated in literature. Data mining process is composed of some set of structured steps that makes the data mining research process and methodology. Initially understanding and analysis of the situation and business problem is completed which is followed with the examination and pre-processing of data. Later a conceptual framework or model is devised following the literature review and analysis approaches. Testing of the model with supervised and unsupervised versions of machine learning approaches takes place. Finally predicted analysis results are evaluated and assessed (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

It is known that data mining is one of the important research methodologies applied in todays contemporary world in order to gain insights and knowledge associated with several phenomena. As Özerk claims, many data mining processes today apply a technical approach in supervised learning where independent or multivariate indicators and variables are assigned to output class labels using functions of mapping. In unsupervised versions of data mining and machine learning, the core values of each cluster (centroids) are calculated, the corresponding sample and attribute values are assigned to the respective clusters so as to maximize convergence and minimize differences in the sam cluster, whereas a divergence is expected with the members of different clusters. In the process of supervised and unsupervised machine learning, rules are created to improve the

exploratory and confirmatory understanding of the phenomenon (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018). In this context, an Aristotelian research design path can bring several advantages in understanding these phenomena and can be a good decision support tool for key business leaders, political leaders and society in general (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

RESEARCH APPROACH

Data mining has been one of the popular research forms that is heavily and frequently used in several sectors and industries as marketing, healthcare, politics, telecommunication, banking and retail. Data mining approach can be considered as a systematical, structured research process which focuses on situation analysis, data gathering, model formation and testing of the model. Later findings and knowledge discovered from these analyses can be used as a decision support point for leaders, science community and society at large. Machine learning technique which is a famous approach in data mining based quantitative research methodologies is a form of learning in machine forms. This learning process is usually triggered and activated by forward feeding approaches which is later followed with backpropogation processes which are stochastic in nature. With the help of mapping functions input layers in the model are mapped to the output layer considering the independent, dependent values. Functions and equations involved in this mapping are calculated. Later in many forms rules generated with the least error rate and which provides the most proximity to actual results are selected and presented as the distinctive association rules. In the evaluation of this a stochastic backpropagation technique is used in many cases. Supervised learning and unsupervised learning are two forms of machine learning in data mining. Supervised learning is a form of classification approach where input and output layer mappings are done with the transformation functions, with the aim of rule discovery and insights discovery. A stochastic backpropagation technique is used in this type of machine learninge in general. In the mapping process independent multivariate variables are assigned to the respective class labels which are considered as the dependent variables in the output layer. Since the initial labeling of the dependent values in the form of nominal values technique, this type of machine learning is named supervised machine learning. On the other hand in unsupervised machine learning several attributes of different instances are assigned to respective clusters with respective values without the requirement of an initial class label declaration. In this form of machine learning several mathematical and statistical functions utilizing heuristics in many cases are applied. In most of the clustering analysis which is also known as unsupervised machine learning centroid values for each independent cluster is calculated and related attribute values for several instances in the data set are assigned to the cluster with the focus on similarity and convergence maximization in one cluster and divergence, difference maximization with other cluster members having other centroid values. In deep learning which is a form of machine learning, feature extraction and classification is integrated in labeling the association rules. In contrast to traditional machine learning algorithms, in deep learning new associatons and features are formed and discovered based on early input values in the latent neurons which is followed with a classification approach in an integrated way. Following the feature extraction and classification steps association rules are assigned to the respective class labels in the model training later the model is tested as in the conventional machine learning approaches. Deep learning can be in the form of supervised, unsupervised or semi supervised fashion. Multilayer Perceptron, Bayesian Networks, Dl4jMlpClassifier (Deep Learning), OneR Method, Hoeffding Tree, Random, Tree, Kmeans have been some of the mostly cited supervised and unsupervised machine learning techniques which utilizes different classification and clustering approaches in literature . Machine learning utilization of data mining can provide exploratory and confirmatory understanding in the phenomena in question and may provide insights and in-depth understanding with knowledge discovery, prediction or forecasting option it provides. In this context a data mining approach strengthened with deep learning techniques have been employed to understand phenomena (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).
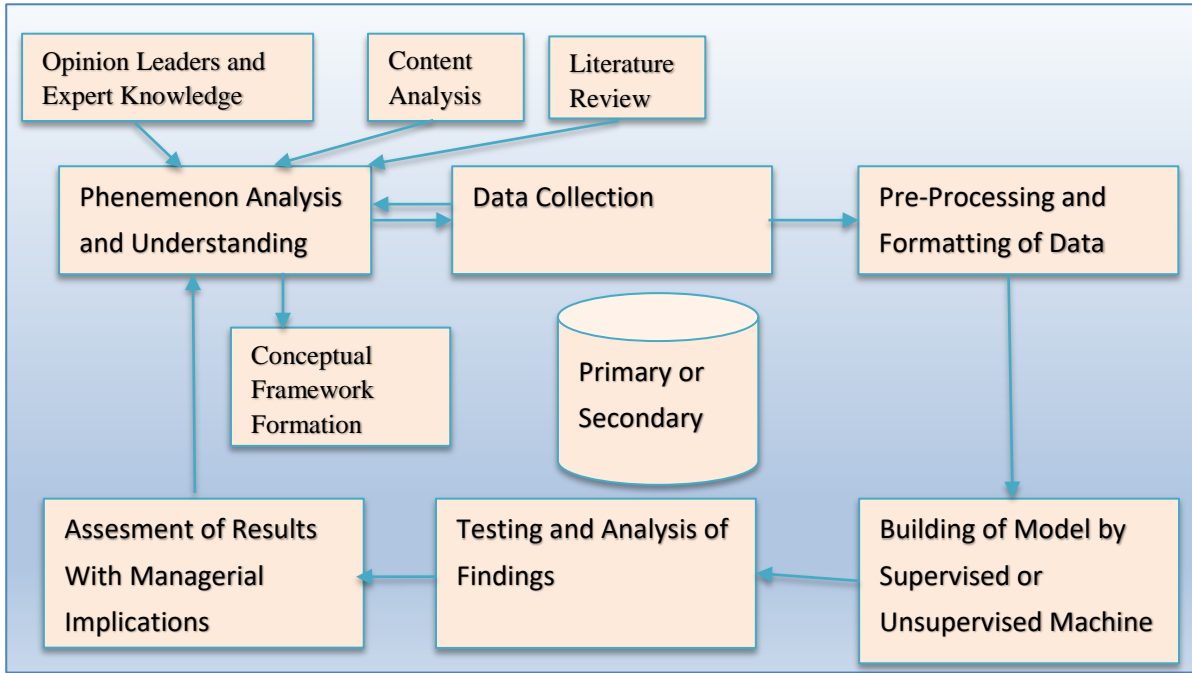
Figure 1. Data Mining Process (Prepared by the Researcher)

In the analysis part, Multilayer Perceptron, Bayesian Networks, Dl4jMlpClassifier (Deep Learning), OneR Method, Hoeffding Tree, Random, Tree, Kmeans were used for machine learning techniques. . The unsupervised machine learning algorithms here evaluate sample values and assign these individual values to sets of relevant segments, while the supervised machine learning algorithms mainly focus on mapping multivariate variables in the input layers to output class labels using transform and mapping functions which is followed by stochastic backpropagation techniques in many cases. In the analysis part the performances of various machine learning approaches are compared and rules with rule functions are built in a reinforced manner, some of them applying forward and backward propagation approaches, depending on their algorithmic architectures and designs (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018). Depending on various factors, such as algorithmic design, algorithmic architecture, algorithmic complexity, these algorithms can produce different results for similar, same or different problem sets and domains (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018). The performance indicators of the algorithms were evaluated and evaluated using the same parameter values with the same data set. With the analyzes performed, the algorithm with the top performer score was discovered with the same data set and parameters maintained. The information patterns and the rules found are listed following the interpretation phase of the research.
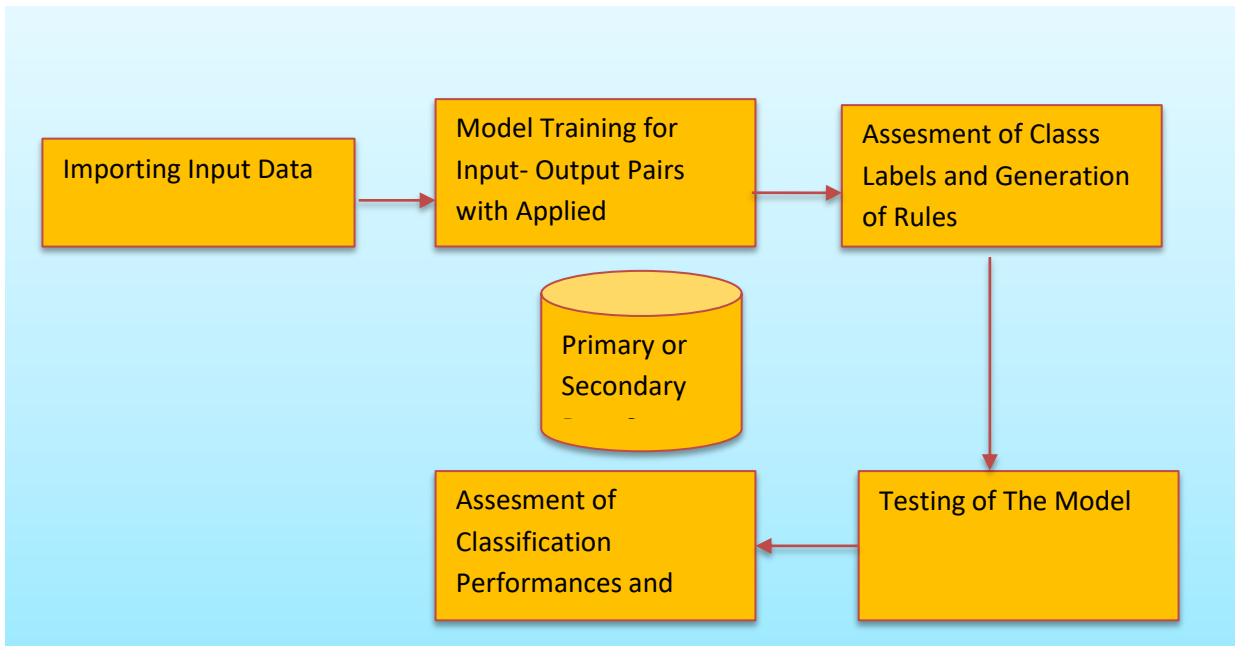
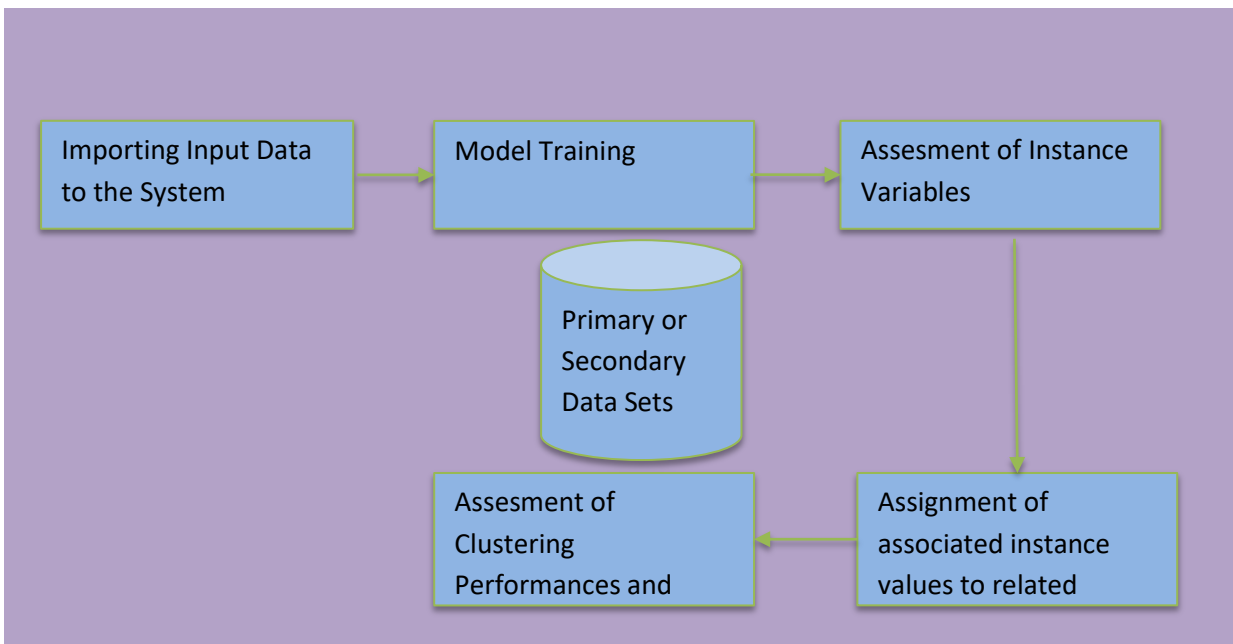Figure 2. Supervised Machine Learning Process (Prepared by the Researcher)

Figure 3. Unsupervised Machine Learning Process Composed of Model Building and Testing (Prepared by the Researcher)

For data set an online administred survey has been conducted using a snowball sampling process (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

Table 1. List of Attributes

Citizenship Status	Nominal
Age Range	Nominal
Marital Status	Nominal
Sexual Orientation	Nominal
Education Status	Nominal
Income	Nominal

ANALYSIS APPROACH

In data mining analysis, association rules, knowledge and understandings are discovered with the help of classification and clustering algorithms for the relevant problem set and domain. In these approaches, input-output mapping functions are used to create association rules that map the outer layer to the inner layer. In some, feed forward and back propagation techniques have been applied. The relevant rules with the least error rate are presented as the main rules of the analysis (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

The same input load with the same parameters was tested using machine learning algorithms, Multilayer Perceptron, Bayesian Networks, Dl4jMlpClassifier (Deep Learning), OneR Method, Hoeffding Tree, Random, Tree, Kmeans. The University of Waikato's Weka data mining package, which includes supervised and unsupervised machine learning applications, was used in the analysis. Then, the performance of classification and clustering was compared and evaluated. In the analysis, 10 fold cross validation method has been used to train and test the model. Based on the performance indicators associated with the data mining analysis, a high performance algorithm was chosen and can be used for such areas and sets of problems to gain additional insight and insight. For this purpose, values of mean squared error, precision, correct classification rate and misclassification rate were used (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).The analysis revealed the performance indicator values and rules as in Tables 2 and 3.

Table 2**.** Performance Scores of Machine Learning Algorithms

Method Applied / Performance Indicator	Dl4jMlpClassifier (Deep Learning)	Naive Bayes	J48	Random Tree	Multilayer Perceptron
RMSE	0.43	35.55	0.44	0.44	**0.41**
Correctly Classified %	44.44	64.44	55.55	68.88	**68.88**
Incorrectly Classified %	55.55	0.42	44.44	31.11	**31.11**

Table 3. Association Rules Generated by machine learning algorithms in data mining

If income is 5000-10000 TL or below 5000 then female, if income is above 10000 TL then male
Incomes in female groups lower compared to the male groups
If gender is female, education status is university graduate or no university graduate then income is in the 5000-10000 TL range
If gender is male and education status is university graduate then income is in the range of 5000-10000 TL
If gender is male and education status is no-university graduate then income is below 5000
If income is in the range of 5000-10000 TL and gender is female then education status is university graduate, if income is in the range of 5000-10000 TL and gender is male then education status is university graduate.
If income is below 5000 TL and gender is female then education status is university graduate, if income is below 5000 and gender is male then education status is university graduate.
If income is above 10000 TL and gender is female then education status is no university degre, if income is above 10000 and gender is male then education status is university graduate.
If age range is between 25-60 and 60 or above then individual is married, if age range is 18-25 then single
If income is below 5000 then single, if income is in the range of 5000-10000 then married, if income is above 10000 then married.

If marital status is married, single or divorced then income is in the range of 5000-10000, if marital status is having a relationship then income is below 5000
If Turkish citizen or non-Turkish citizen then income is in the range of 5000-10000
If sexual orientation is heterosexual or other then income is in the range of 5000-10000 whereas if sexual orientation is Bisexual then income is below 5000

The analysis results revealed that if income is 5000-10000 TL or below 5000 then female, if income is above 10000 TL then male. Incomes in female groups lower compared to the male groups. If gender is female, education status is university graduate or no university graduate then income is in the 5000-10000 TL range. If gender is male and education status is university graduate then income is in the range of 5000-10000 TL. If gender is male and education status is no-university graduate then income is below 5000. If income is in the range of 5000-10000 TL and gender is female then education status is university graduate, if income is in the range of 5000-10000 TL and gender is male then education status is university graduate. If income is below 5000 TL and gender is female then education status is university graduate, if income is below 5000 and gender is male then education status is university graduate. If income is above 10000 TL and gender is female then education status is no university degre, if income is above 10000 and gender is male then education status is university graduate. If age range is between 25-60 and 60 or above then individual is married, if age range is 18-25 then single. If income is below 5000 then single, if income is in the range of 5000-10000 then married, if income is above 10000 then married. If marital status is married, single or divorced then income is in the range of 5000-10000, if marital status is having a relationship then income is below 5000. If Turkish citizen or non-Turkish citizen then income is in the range of 5000-10000. If sexual orientation is heterosexual or other then income is in the range of 5000-10000 whereas if sexual orientation is bisexual then income is below 5000
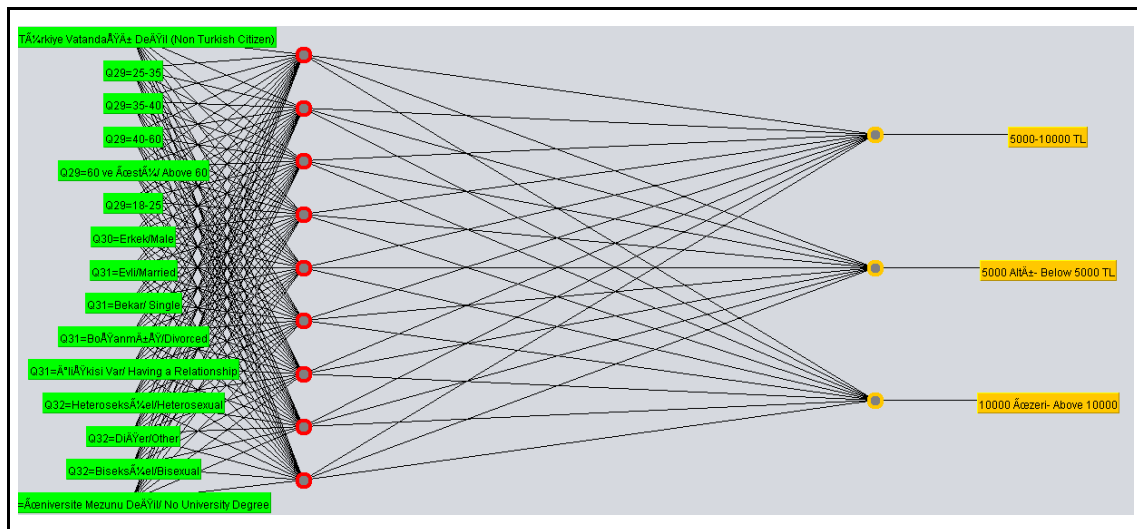

Figure 4. A Neural Network View of the Model Generated (Multi-Layer Perceptron- Income is the dependent variable)

Among several methods applied, Multilayer Perceptron method has been the top performing algorithm with a RMSE, correct classification and incorrect classification rates among other supervised machine learning approaches applied. To sum up, supervised and unsupervised machine learning algorithms, which are also known as classification and clustering techniques in data mining literature can be used as an effective and efficient tool for knowledge discovery or confirmation in exploratory and confirmatory research designs. These insights may be considered by decision makers and society at large in such problem sets and domains. Based on the input loads, algorithmic design, architecture and performance of the algorithm which can be assessed with approximations, metrics as Big O or Big Ω which are used to assess the efficiency and the computational complexity (Özerk, 2021, Özerk, 2022, Özerk, 2009, Özerk, 2018).

CONCLUSION

Throughout the history several genders, races, ethnic groups, workers, individuals with several sexual preferences and orientations, individuals with different colours, life styles and preferences, freedom of speech and thought, in a broader range many human beings, animals qualified accaptence by different members of the society and won their rights in the legal framework and constitutions by law. Human rights, animal rights, worker rights, women rights, equality of women and men, right to elect and to be elected, freedom of speech and thought are some of these rights and developments human race experienced and saw during time as a result. However as in the past it is misfortune in some of these geographies and business settings still some forms of discrimination, stigmatization and problems alike occur. As the paradigms of the society, change and new developments enter into our lives we see many changes and developments in our lives which are strengthened with rules, regulations and legislative frameworks. However, unfortunately there is still lack of equal access to resources, double standardizations, inequalities of opportunities with several inequalities in today's world. In this research gender pay gap, which is frequently cited as one of the important problems in today's world is examined with potential influences associated with it using deep learning, machine learning and data mining methodologies.

Today inequalities in business world, inequalities in access to resources, inequalities in pays, gender, color and nationality based discriminations and double standards remain to be some of the problems that is seen in some countries, geographies and business settings. As stated in literature, gender based income inequalities in business landscape remains to be an important problem in several business settings, countries and geographies which leads to several negative consequences in 21st century. As in other lower income segments, lower income female groups are facing difficulties in affording their living expenses especially in key areas as, education, health, nutrition, housing, self development. They face difficulties in sustaining their lives and the generations that they raise arising from financial difficulties compared to higher income groups of the society. Since they do not afford the education for theirselves and their children, these generations can not have the necessary education to develop capacities and capabilities to be successful and take place in the job market.

The analysis results revealed that if income is 5000-10000 TL or below 5000 then female, if income is above 10000 TL then male. Incomes in female groups lower compared to the male groups. If gender is female, education status is university graduate or no university graduate then income is in the 5000-10000 TL range. If gender is male and education status is university graduate then income is in the range of 5000-10000 TL. If gender is male and education status is no-university graduate then income is below 5000. If income is in the range of 5000-10000 TL and gender is female then education status is university graduate, if income is in the range of 5000-10000 TL and gender is male then education status is university graduate. If income is below 5000 TL and gender is female then education status is university graduate, if income is below 5000 and gender is male then education status is university graduate. If income is above 10000 TL and gender is female then education status is no university degre, if income is above 10000 and gender is male then education status is university graduate. If age range is between 25-60 and 60 or above then individual is married, if age range is 18-25 then single. If income is below 5000 then single, if income is in the range of 5000-10000 then married, if income is above 10000 then married. If marital status is married, single or divorced then income is in the range of 5000-10000, if marital status is having a relationship then income is below 5000. If Turkish citizen or non-Turkish citizen then income is in the range of 5000-10000. If sexual orientation is heterosexual or other then income is in the range of 5000-10000 whereas if sexual orientation is bisexual then income is below 5000

Among several methods applied, Multilayer Perceptron method has been the top performing algorithm with a RMSE, correct classification and incorrect classification rates among other supervised machine learning approaches applied. To sum up, supervised and unsupervised machine learning algorithms, which are also known as classification and clustering techniques in data mining literature can be used as effective and efficient technique for knowledge discovery or confirmation in exploratory and confirmatory research designs. These insights may be considered by decision makers and society at large in such problem sets and domains with the exploratory and confirmatory understandings they provide.

Biography of Author

Asst. Prof. Dr. Özerk Yavuz received his Ph.D. degree in Business Administration-Marketing from Bahcesehir University, Istanbul, M.Sc. degree in Computer Engineering from Bahcesehir University, Istanbul and his B.Sc. degree in Computer Technology and Information Systems from Bilkent University, Ankara. Several papers and articles of him have been published in respected and prestigious refereed, international scientific journals, books, book chapters, conference proceedings and presented in international conferences and congresses. Dr. Özerk Yavuz also has been referee, reviewer, moderator or editor of several notable, trusted international scientific journals and international, scientific, academic books. He is interested in management information systems, software engineering, computer engineering, data mining, virtual communities, virtual networks, marketing, management, and business administration. Dr. Özerk Yavuz has abroad and domestic working experiences in several institutions and countries, in various fields of business and higher education. He is interested in Salsa, Rumba, Cha-cha, East Coast Swing, Argentine Tango, American Tango, Vienna Waltz, Milonga and has been an active member of Bilkent University dance community. In his free time, he loves travelling, swimming and enjoying different kitchens. Dr. Özerk Yavuz has worked with several respected and distinguished scholars, leaders and teenagers in his work life. He has been a member of several distinguished scientific communities, Bilkent University and Bahçeşehir University alumni organizations. He is currently working in Halic University, Faculty of Management, Management Information Systems department as Asst. Prof. Dr. and continues his academic, administrative works.

Email: dr.ozerk@gmail.com; ozerkyavuz@halic.edu.tr;ozerk@alumni.bilkent.edu.tr

Personal Web Site: https://sites.google.com/view/drozerkyavuz

Orchid: https://orcid.org/0000-0002-1371-688X

Scopus: https://www.scopus.com/authid/detail.uri?authorId=55371850200

Google Scholar: https://scholar.google.com/citations?user=hR7QlJMAAAAJ&hl=en

Publons: https://publons.com/researcher/4754966/dr-ozerk-yavuz/

Researchgate: https://www.researchgate.net/profile/Oezerk-Yavuz

REFERENCES

Kotler, P. (1991), Principles of Marketing, Englewood Cliffs - Prentice Hall, New Jersey (US).

Kotler, P. (2000), Marketing Management, Upper Saddle River - Prentice Hall, New Jersey (US).

Malhotra, Naresh K. (2007), Marketing Research: An Applied Orientation, Upper Saddle River - Pearson/Prentice Hall, New Jersey (US).

Yavuz, Ö. (2009), "An Early Warning System Approach for the Identification of Currency Crises", Master Thesis, Bahcesehir University Graduate School of Sciences, Istanbul.

Yavuz, Ö. (2018), "Marketing Implications of Participative Behavior in Virtual Communities", Doctoral Thesis, Bahcesehir University Graduate School of Social Sciences, Istanbul.

Yavuz, Ö. (2021), A Data Mining Analysis of Coronavirus Cases and Vaccinations in The City of London, Astana Publisher, Ankara.

Yavuz, Ö. (2021), "A Public Perceptions Analysis with Data Mining Algorithms", 2. International "Başkent" Congress on Physical, Social and Health Sciences, 25-26 September – Ankara, pp.69-73.

Yavuz, Ö., Karahoca, A. and Karahoca, D. (2019), "A Data Mining Approach for Desire and Intention to Participate in Virtual Communities", International Journal of Electrical and Computer Engineering, S.9(5), pp.3714-3719.

Yavuz, Ö. (2021), "A Leading Indicator Approach with Data Mining Techniques in Analysing Bitcoin Market Value." European Journal of Science and Technology.

Yavuz, Ö. (2022), "An Optimization Focused Machine Learning Approach in Analysing Arts Participative Behavior with Fine Arts Education Considerations." International Scientific and Vocational Studies Journal.

Yavuz, Ö. (2022). "A 21st Century Approach in Analysing Health Precautions in London with Machine Learning Driven Data Mining." European Journal of Science and Technology.

Yavuz, Ö. (2022). "Nitel ve Nicel Araştırma Tasarımı Uygulamasıyla Etanol Kullanımını ve Etkilerinin Anlaşılması." European Journal of Science and Technology.

Yavuz, Ö. (2022). "A Classification and Clustering Approach Using Data Mining Techniques in Analysing Gastrointestinal Tract." International Scientific and Vocational Studies Journal.

Yavuz, Özerk. (2022). Engıneerıng & archıtectural Scİence - 2 (A Multdıscipliner Approach).

Karahoca, D., Karahoca, A. and Yavuz, Ö. (2013), "An Early Warning System Approach for the Identification of Currency Crises with Data Mining Techniques", Neural Computing and Applications, S.23(7-8), pp.2471-2479.