Kafkas Univ Vet Fak Derg
23 (4): 673-680, 2017
DOI: 10.9775/kvfd.2016.17281

Kafkas Universitesi Veteriner Fakultesi Dergisi
Journal Home-Page: http://vetdergi.kafkas.edu.tr
Online Submission: http://vetdergikafkas.org

Review

A Review of Machine Learning Applications in Veterinary Field

Pınar CİHAN [1] Erhan GÖKÇE [2] Oya KALIPSIZ [1]

[1] Department of Computer Engineering, Faculty of Electrical Electronic, University of Yıldız Technical, TR- 34220 Esenler, Istanbul - TURKEY
[2] Department of Internal Medicine, Faculty of Veterinary Medicine, University of Kafkas, TR-36100 Kars -TURKEY

Article Code: KVFD-2016-17281 Received: 23.12.2016 Accepted: 24.03.2017 Published Online: 27.03.2017

Citation of This Article

Cihan P, Gökçe E, Kalıpsız O: A review of machine learning applications in veterinary field. *Kafkas Univ Vet Fak Derg,* 23 (4): 673-680, 2017. DOI: 10.9775/kvfd.2016.17281

Abstract

Machine learning is a sub field of artificial intelligence which allows forecasting through learning past behaviors and rules from old data. In today's world, machine learning is being used almost in any fields such as education, medicine, veterinary, banking, telecommunication, security, and bio-medical sciences. In human health, although machine learning is generally preferred particularly in predicting diseases and identifying respective risk factors, it is obvious that there are a limited number of publications where this method was applied on veterinary or indicates whether it is correct and applicable. In this review, it was observed that the neural network, logistic regression, linear regression, multiple regression, principle component analysis and k-means methods were frequently used in examined publications and machine learning application in veterinary field upward momentum. Additionally, it was observed that recent developments in the field of machine learning (deep learning, ensemble learning, voice recognition, emotion recognition, etc.) is still new in the field of veterinary. In this review, publications are examined under clustering, classification, regression, multivariate data analysis and image processing topics. This review aims at providing basic information on machine learning and to increase the number of multidisciplinary publications on computer sciences/engineering and veterinary field.

Keywords: *Machine learning, Artificial intelligence, Veterinary, Computer sience*

Veteriner Hekimlik Alanında Makine Öğrenmesi Uygulamaları Üzerine Bir Derleme

Özet

Makine öğrenmesi yapay zekanın bir alt çalışma alanı olup eski verilerden geçmiş davranışların ve kuralların öğrenilerek ileriye doğru tahminlerin yapılmasına olanak sağlar. Makine öğrenmesi günümüzde eğitim, tıp, veterinerlik, bankacılık, telekomünikasyon, güvenlik, biyomedikal gibi hemen hemen her alanda kullanılmaktadır. İnsan sağlığında özellikle hastalıkların önceden tahmin edilmesi ve ilgili risk faktörlerinin tespit edilmesinde makine öğrenmesi yöntemleri genellikle tercih edilmekle birlikte hayvan sağlığında doğruluğu ve aynı zamanda ilgili alanda kullanılabilir olup olmadığını belirleyen bu yöntemin uygulandığı çalışmaların sınırlı olduğu görülmektedir. Bu derlemede incelenen çalışmalarda sinir ağları, lojistik regresyon, lineer regresyon, çoklu regresyon, temel bileşen analizi ve k-ortalamalar yöntemlerinden sıklıkla yararlanıldığı ve veterinerlik alanında yapılan makine öğrenmesi çalışmalarının son yıllarda ivme kazandığı gözlemlenmiştir. Ayrıca makine öğrenmesi alanındaki son gelişmelerin (derin öğrenme, kolektif öğrenme, ses tanıma, duygu tanıma, vb.) veterinerlikte yeni yeni uygulandığı gözlemlenmiştir. Bu derlemede çalışmalar kümeleme, sınıflandırma, regresyon, çok değişkenli veri analizi ve görüntü işleme başlıkları altında incelenmiştir. Bu derlemenin amacı makine öğrenmesi ile ilgili temel bilgileri vermek ve bilgisayar bilimleri/mühendisliği ile veterinerlik alanındaki ortak çalışmaları arttırmaktır.

Anahtar sözcükler: *Makine öğrenmesi, Yapay zeka, Veterinerlik, Bilgisayar bilimleri*

INTRODUCTION

Machine learning deals with data analysis which aims to build analytical models in an automated way. Machine learning helps computers to discover insight information by learning from data through iterative algorithms, even if they are not programmed to search for them [1].

Thanks to machine learning methods, the machines have become capable of contributing to the brain power of the humanity as well as their contribution to manpower. These methods support us in creating assumptions on the future by analyzing a large amount of data for any practice and they help us in decision making. Therefore, the importance and contribution of machine learning methods are increasing more and more each day [2].

As new detection and diagnostic modalities are developed and data types getting complexer and multi-modal analysis getting more important which causes a

İletişim (Correspondence)
+90 212 3835752
pinar@ce.yildiz.edu.tr

remarkable increase of collected data in livestock, it is obvious that machine learning is required more than ever with great potential for future use. Multi-dimensional and complex datasets confronted by researcher may be interpreted through new tools enabled by machine learning.

There is a close relation between machine learning and fields such as artificial intelligence, data mining, statistics and especially probability theory, pattern recognition and computer sciences in general. Undoubtedly, there are many applications currently used in fields such as health and medicine [3], drug design [2], banking [4,5], education [6,7], telecommunications [8], software development [9], bio-medicine, security, geology, astronomy. Other fields of use are fingerprint recognition [10], iris recognition [11], face recognition [12], handwriting and signature recognition [13], medical data identification [14], internet search engines [15] etc. *(Fig. 1)*.

In medicine, especially computer aided disease diagnosis and exploration of risk factors and their relations to diseases, very good results have been achieved by machine learning techniques. Machine learning techniques are distinguished with their successful results in relation to human health and animal health studies and they may provide new frontiers to solve problems which are known to be persistent in the field.

In several medical studies which are important for human health such as breast cancer prediction [16], estimation of cancer types [17], estimation of survival in patients with severe burns [18], identification of risk factors that trigger heart attacks [19], early detection of cancer [20], heart and vascular disease diagnosis [21], anomaly detection [22], polysomnographic [1], machine learning methods were benefited from.

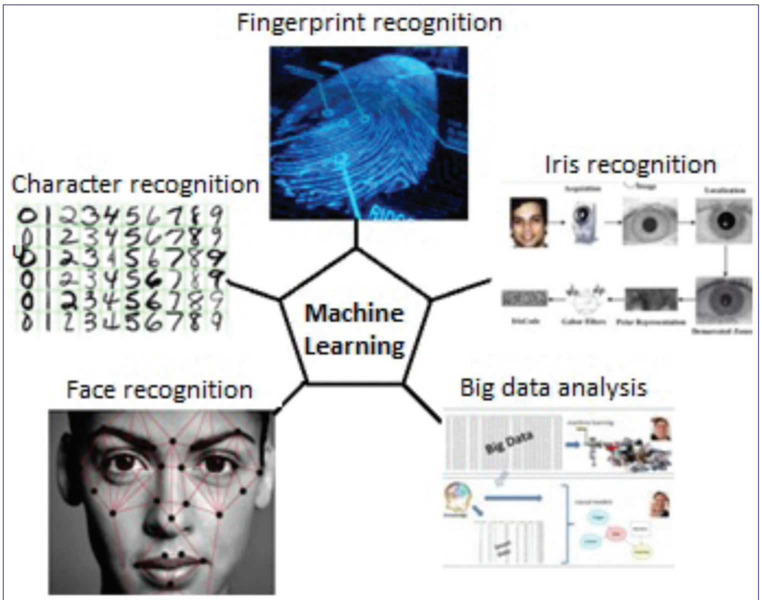

Fig 1. Application fields of machine learning

Machine learning is obviously one of the areas where most of researches were conducted recently. Thus, the importance of exploring this area is increasing. For analyzing various aspects of livestock farming which is closely related to human health since the very early days of history, such as the estimation and determination of animal diseases and respective risk factors, machine learning applications are required. Machine learning on veterinary is a new application method and it has drawn some attention recently and some studies were conducted in this area [23-27].

In this review, machine learning methods were explained briefly, in order to foster the implementation of machine learning methods in the field of veterinary and to provide basic information to researchers. Applications of clustering, classification, regression, feature analysis and image processing on veterinary were investigated. We suggest that the researchers may acquire a basic idea on machine learning through this study, so that they refer to machine learning methods for solving problems on animal health. In this way, the number of machine learning applications on veterinary shall increase.

MACHINE LEARNING TECHNIQUES

Machine learning is a way of programming computers where a performance criterion is optimized thanks to recent data or past experience. We define a model with some variables and through training data or past experience a computer program is taught to optimize the parameters of this model. The model may either be designed as a predictive one to predict the future, or a descriptive model to acquire knowledge from the dataset, or both [28]. *Fig. 2* outlines one way in the steps might be incorporated into an end to end machine learning system for analyzing data from a science or engineering application.

Some advantages of the machine learning are as follows [29]:

- In some cases, whenever we specify the input/output it may be impossible to estimate the relations among various data and to indicate their relationship. Thus, the machine is expected to estimate their relationship by setting its own microstructure.

- If the amount of information is too much it is likely to be handled by machines rather than humans.

- The machine is able to adapt to change easily.

- Machine learning methods reveal concealed relation-ships and links through a large data stack.

Fig. 3 shows the process of the machine

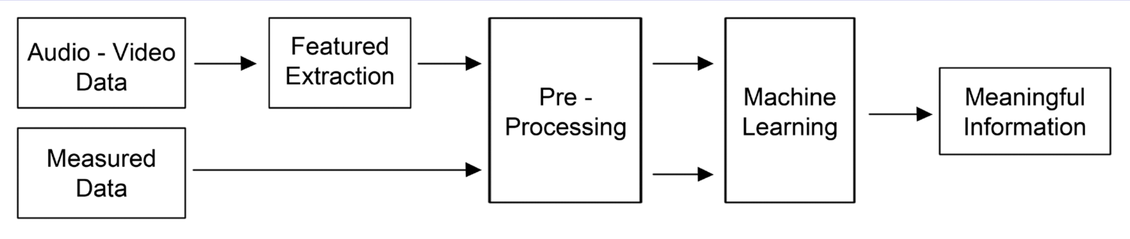

Fig 2. General piplines of machine learning

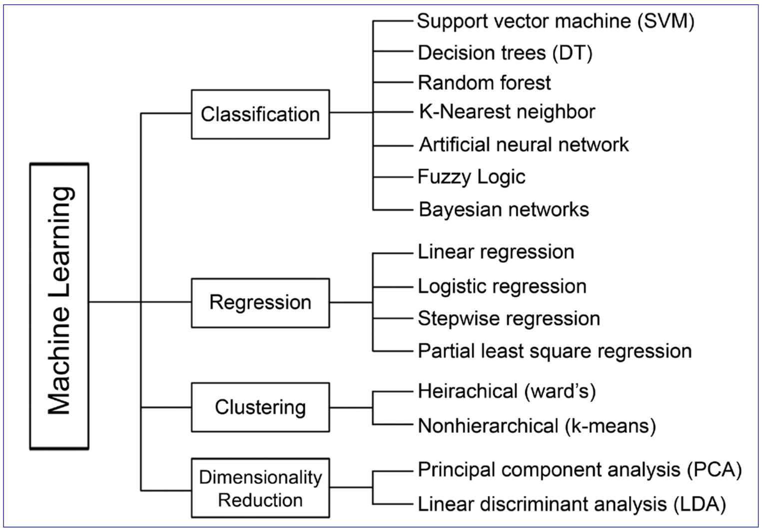

Fig 3. Techniques of machine learning and commonly used algorithms

learning and commonly used algorithms in computer science field. Techniques of machine learning model as categorized into supervised and unsupervised methods. The supervised machine learning techniques uses algorithms to reason from external instances which then generate universal theorems in order to predict future instances. This type of machine learning constructs a brief model for allocation of class labels to predict features. The resulting classifier is used to assign class labels to testing instances, the significance of the predicting feature is known, but there is no data about the value of the class label. The unsupervised methods do not utilize surrounding data to generate objective output or benefits.

Classification is a process to establish a model which is used to describe and discriminate between data classes. A group of training data is analyzed to derive a model. For objects with unknown class labels, the categorical (discrete, unordered) class label is predicted by this model. The process of data classification consists of a learning step and a classification step. A classifier is generated at the first step, i.e. the learning step (or training phase) which describes a preset group of data classes. We test the accuracy of this model at the second step and if it is positive, we use it for classifying the new data [6].

Regression analysis is the most frequently used among statistical methodologies for numeric prediction. While categorical labels are predicted through classification, continuously valued functions are modeled by regression. Rather than predicting (discrete) class label, regression is used for missing or unavailable numerical values [30,31].

Clustering analyzes data objects without consulting class label, unlike classification and regression, which are utilized in analysis of class labeled (training) datasets. Generally, it is quite usual for class labeled data to be missing at the early stage. Class labels may be created for a dataset through clustering. By maximizing and minimizing the interclass similarity the objects may be grouped or clustered. Clustering of objects take place in order to compare the similarity of objects within a cluster to one another, but without similarities to objects in other clusters [2].

Clustering methods are divided in two categories. These are hierarchical clustering and non-hierarchical clustering, also known as partitioning. The hierarchical methods (often known as k-means clustering methods) produce a set of nested clusters in which each pair of objects or clusters is progressively nested in a larger cluster until only one cluster remains. The non-hierarchical methods divide a dataset of *N* objects into *M* clusters, with or without overlap. Each object is a member of the cluster with which it is most similar, however the threshold of similarity has to be defined [32].

Image processing is used as a method to work on an image, in order to enhance it or to find out some useful information about it. Image processing grows rapidly today with its many application areas. It is established as a core research area within fields of engineering and computer sciences. Image processing uses two methods. These are analogue and digital image processing methods. For hard copies such as printed materials and photographs analogue image processing is used. Image analysts implement many interpretation techniques when these methods are used [33].

Multivariate data analysis is a statistical method for

***Table 1.** Machine learning applications on veterinary field*

Classification Studies				
Author	**Year**	**Method**	**Objective**	**Results**
Akıllı et al.[35]	2016	Fuzzy Logic	They have designed a decision support system based on fuzzy logic and tried to determine the compatibility between the system and expert decision.	They have reported that the decision support system designed using the records on reproduction and milk production efficiency of Holstein Friesian dairy cows was conducting an accurate classification with a success rate of 92.6% and that fuzzy logic based decision support systems shall be successful on livestock farming.
Hempstalk et al.[36]	2015	DT, NB, BN, LogR, SVM, PLSR, RF, RotF	They have attempted to estimate the success of insemination and conception in dairy cows.	In this study, they have used 8 different machine learning methods and they have reported that the logistic regression method has had the best performance in general.
Küçükönder et al.[24]	2014	ANN, RBF Network, NB, KStar, Ridor	They have studied the classification on fertilization of Japanese quail eggs according to factors such as season, natural selection and frequency of settlement and to determine the influences of these factors.	In this study, 85% of quail eggs were determined to be fertile and 15% of them to have lower reproduction capacities with a success rate of 99.73%. It was observed that the Ridor algorithm for classification of eggs according to their fertility (as fertile or unfertile) has generated better results with lower error rate and the fertility rate was determined as 85.71%.
Lewis et al.[37]	2011	BN	They have tried to indicate that Bayesian network is an analytical method for complex animal health data.	As a result of this study they have reported that the statistical inference of Bayesian network model offers a richer analytical tool in comparison to any standard statistical technique.
Karabag et al.[38]	2009	DT	They have tried to determine the influential factors on hatching ability for Chucar partridges eggs.	They have determined overall hatching, fertility and hatching ability as 56.2%, 79.2%, and 71.0% respectively and they have reported that the classification tree method has predicted the external egg traits such as, egg weight, egg volume, egg length and width was a significant factor on hatching ability with an accuracy of 75.6%.
Pelaez and Pfeiffer[39]	2008	LogR, DT, FA	They have tried to classify cattle herds according to the presence of infectious disease.	In this study, they have reported that there was high risk in regions where the cattle population is dense and in many regions of Wales which are closed and where the cattle movement is frequent.
Regressional Studies				
Author	**Year**	**Method**	**Objective**	**Results**
Gokçe et al.[31]	2014	Simple/ multiple regression	They have tried to analyze the relationship between serum lactoferrin concentrations and serum IgG concentrations in lambs.	They have reported that there was a significant linear correlation (R^2=0.375) between serum lactoferrin concentrations as a predictor of passive immunity and serum IgG concentrations in lambs during different days of neonatal period (1st, 2nd, 4th, 7th, 14th and 28th day), but that it was insufficient for calculating the IgG concentration.
Gokçe et al.[40]	2013	Chi-square, Odds, RR	They have tried to investigate the influence of some factors on diseases and death of sheep at neonatal period and afterwards.	They have reported that the most important risk factors of lamb diseases and deaths were lambing season, number of births per dam, birth weight and dam's health status.
Gokçe et al.[41]	2013	MSRA, Simple/ multiple regression	They have tried to determine the relationship between passive immunity and growth performance in lambs during and after the neonatal period and to assess the impact of some factors on the growth performance.	They have reported that passive immunity in lambs has varied significantly in the early 12 weeks, that growth performance was depending on the birth weight, type of birth, gender, healthiness, dam's age and lambing season by indicating that growth performance was reduced if the dam's age was ≤2 years or it was a twin birth or a female lamb was born or the lamb was born in the winter or it has got ill or if the birth weight ≤3 kg.
Teke et al.[42]	2013	LinR, MLP, SMOreg	They have tried to model the live weights of Holstein Friesian breed using their body sizes.	They have reported that Linear Regression model, Multilayer Perceptron and SMOreg had success rates of 97.94%, 97.72% and 99.17% respectively and that it was possible to predict live weights through data mining with a high reliability.
Ghotoorlar et al.[25]	2012	ANN	They have tried to develop an automatic scoring system complying with the subjectively calculated lameness score.	23 features of 105 freely moving dairy cows were used and the cows were divided into 5 groups according to their movement score. They have reported that group No.1 and No.4 had the highest sensitivity and specificity value.
Takma et al.[43]	2012	Multiple regression, ANN	They have tried to research the effects of duration of lactation, calving year and service period in cows on lactation milk yield and their adaptability.	They have reported that the ANN model was more appropriate than multiple regression model in estimation of milk yield of Holstein Friesian cows and that it might be an alternative method to regression analysis, because it produces results with fewer errors.

***Table 1**. Machine learning applications on veterinary field (Continue...)*

Author	Year	Method	Objective	Results
Clustering Studies				
Author	**Year**	**Method**	**Objective**	**Results**
Bank et al.[44]	2015	K-means	They have tried to characterize some gut bacteria which were found in piglets within the first week of their birth.	They have reported that there is a larger variation between the disturbed bacterial composition of neonatal piglets and diarrhoeic piglets and that diarrhoea was affecting the first week of neonatal piglets.
Nantima et al.[45]	2015	Ward's	They have tried to identify the risk factors associated with the occurrence and spread of African swine fever among smallholder pig farmers	They have used to ward's hierarchical clustering method to determine the number of clusters. According to their analyses they have observed some significant differences among the three cluster. for instance, households in cluster 1 that had purchased the least number of pigs reduced their risk to ASF compared to the other two clusters. Cluster 3 that was most vulnerable had the majority of the households practicing free range which increased ASF risk into these farms. Although cluster 2 had more households feeding swill. few of these households were affected by ASF outbreaks.
Dupuy et al.[46]	2013	K-means, Hierarchical clustering, PCA, MFA	They have tried to group the cows using carcass and health-related data.	They have obtained 12 consistent clusters according to slaughtering year and slaughterhouse and they have reported that the combination of their clustering methods with multiple factor analysis were appropriate for larger and complexer slaughterhouse data.
Petit et al.[47]	2010	K-means	They have tried to group the wild animals according to syndromes using wildlife necropsy data.	They have reported that they have obtained 9 clusters and that these clusters reflect the most obvious and frequent diseases. They have also indicated that k-means was a useful tool.
Dogan [48]	2002	K-means	He has tried to demonstrate that clustering analysis method might be applied in studies on animal breeding using some body sizes of the Arabian fillies.	He has reported that the height at withers, heart girth and cannon bone circumference of fillies were lower and that the differences among genders were rather caused by the genetic structure due to the gender than age and he has indicated that clustering analysis would be appropriate especially during selection at animal breeding.
Gürcan et al.[23]	2002	Hierarchical clustering	They have tried to classify the German Meat Merino and Karacabey Merino Sheep genotypes using live weight, body size, and fiber diameter.	After the clustering analysis it was discovered that two genotypes have similar body sizes and for both genotypes it was observed that the subgroup consisting of 1.5 and 2.5 year old sheep indicated heterogeneity, whereas the subgroup consisting from 3.5, 4.5 and 5.5 year old sheep showed homogeneity. Besides they have reported that 98.9% of the herd was found in the same cluster when both genotypes were assessed together and that there were no significant difference between the two in terms of body size.
Multivariate Data Analysis Studies				
Author	**Year**	**Method**	**Objective**	**Results**
Gokçe et al.[26]	2013	MSRA, GLM	They have tried to research some risk factors influencing the passive immunity and birth weight in lambs and to figure out the relationship between the passive immunity and birth weight.	They have reported that some farm administration applications and animal characteristics were related to birth weight and passive immunity and also that birth weight was effective on passive immunity.
Yunusa et al.[49]	2013	PCA	They have tried to analyze the morpho-logical structure of Nigerian Uda and Balami sheep.	They have reported that the most important features for describing both breeds were traits relating to cranial measurements and bone development.
Akçay et al.[50]	2012	PCA	They have tried to evaluate carcass parts of broiler chicken.	10 criteria were used in the study, after the principal component analysis they have determined that the first five principal components were corresponding to the 80.4% of the total variance and that the first principal component was able to describe 42.3% of the total variance. Moreover, they have reported that PCA may be used as a tool in assessing and understanding the total variance in livestock farming
Casanova et al.[51]	2012	PCA	They have tried to classify the cows of French and Spanish breeds according to their morphological properties.	They have concluded that the first principal component was face and skull lengths describing 49.9% of the variance and the second principal component was skull and head width describing 19.2% of the variance.
Meyer [52]	2007	PCA	They have tried to demonstrate that it was possible to reduce the number of properties through PCA using carcass data and ultrasound data of Angus cows.	They have reported that 14 traits were used in the study and that 7 of them sufficed for selection index calculations and estimation of breeding values and that it was possible to halve the number of traits to estimate the breeding values directly through principal components.
Lee et al.[53]	2006	LDA, MFCCs	They have tried to automatically identify animals (30 kinds of frog calls and 19 kinds of cricket calls) from their sounds.	LDA used to reduce the feature dimension and increase the classification accuracy. They have reported that the average classification accuracy is 96.8% for 30 kinds of frog calls and 98.1% for 19 kinds of cricket calls.
Caraviello et al.[54]	2006	DT, BN, Instance-based algorithms	They have tried to determine the variables affecting pregnancy in cows and the variables affecting first-service conception rate.	Through the decision tree algorithm, they have obtained an accurate classification for the pregnancy status at a rate of 71.4% and for the first-service conception a rate of 75.6% in terms of Holstein cows on large dairy farms.
Bilgin and Esenbuga [55]	2005	Canonical correlation analysis	They have tried to evaluate the relationship between the body sizes and and carcass weight of Morkaraman sheep.	They have reported that it is possible to estimate carcass traits of male Morkaraman lambs from their live body sizes and that it may be useful to implement low cost live body measurements in earlier years for carcass weights as a detailed selection criterion.

***Table 1.** Machine learning applications on veterinary field (Continue...)*

Image Processing Studies				
Author	**Year**	**Method**	**Objective**	**Results**
Bozkurt et al.[27]	2013	ANN	They have tried to determine the feeding performance and carcass properties of Brown Swiss and Holstein Friesian breeds in a feedlot system.	They have reported that according to models obtained from digital image analysis and ANN, body length and heart girth were the best predictive variable for estimating the live weight; they have reported that the best predictive variable for estimation of warm carcass weight was the carcass length.
Mcevoy et al.[56]	2013	ANN, PLSDA	They have tried to determine the region including the hip joint on radiographic images of dogs.	They have reported that veterinary images have the potential of being utilized for educational purposes in classification and grouping.
Bilgin et al.[57]	2011	Image Processing Method	They have tried to demonstrate that nose prints of Kangal breed dogs were different from each other.	The resulting values were very different and remote from each other according to statistical data obtained. They have declared that this was caused by the uniqueness of images.
Slósarz et al.[58]	2011	ANN	They have attempted to estimate the fat content within the muscular parts of lambs.	They have reported that there was a significant relation between the body weight and the lamb's age before slaughter, while the relationship between the body weight and intramuscular fat content was weak.

__ANN:__ Artificial neural networks, __BN:__ Bayes networks, __DT:__ Decision tree, __FA:__ Factor analysis, __GLM:__ General linear model, __LDA:__ Linear discriminant analysis, __LogR:__ __Logistic regression__, __MFA:__ Multiple factor analysis, __MFCCs:__ Mel-frequency cepstral coefficients, __MLP:__ Multilayer perceptron, __MSRA:__ Multivariable stepwise regression analysis, __NB:__ Nive bayes, __Odds:__ Odds ratio, __PCA:__ Principal component analysis, __PLSDA:__ Linear partial least squares discriminant analysis, __PLSR:__ Partial least squares regression, __RotF:__ Rotation forest, __RR:__ Relative risk, __SMOreg:__ support vector machine for regression, __SVM:__ Support vector machine

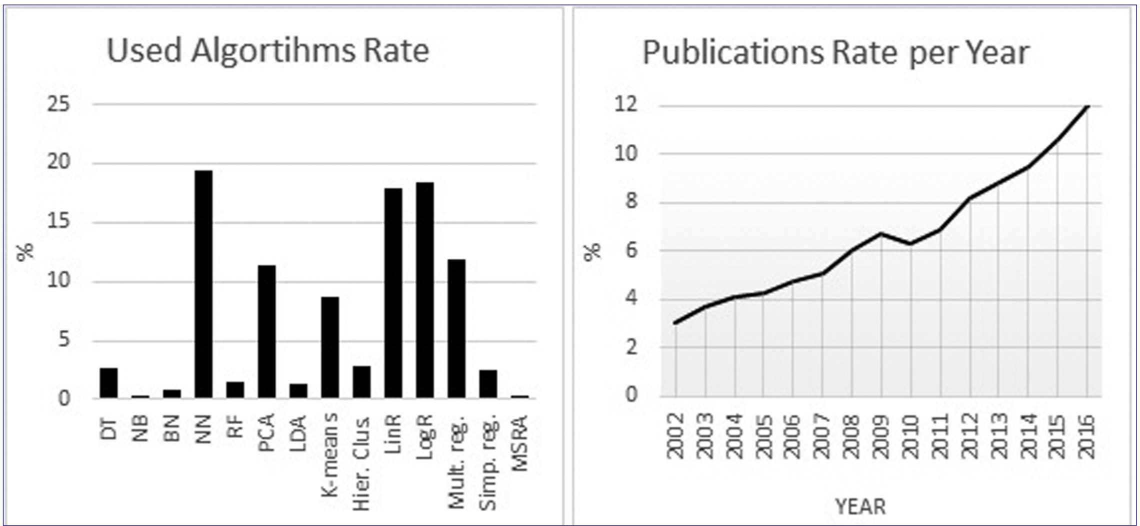

Fig 4. The distrubition of publications according to machine learning methods and publication frequency rate per year

analyzing data consisting of multiple variables. Multiple variables typically come into play in many cases of daily life, products to be chosen or decisions to make and they help modeling the real life. The available data is abundant, but difficulties arise in terms of clarifying the picture and making intelligent decisions based on it [34].

MACHINE LEARNING APPLICATIONS ON VETERINARY

While the articles scanning we selected keywords: "analysis of animal diagnosis", "diagnosis of disease veterinary", "evaluation of risk factor analysis for animal diseases", "animal medical data analysis", "disease decision support animal", "predict animal diagnosis machine learning", "analysis small ruminant diagnosis machine learning", "animal passive immune system machine learning", "naïve bayes in veterinary", "machine learning predict animal disease", "predict animal disease using decision tree", etc. and their combinations are used. The research questions may not cover all machine learning area in detail. We tried to get only a snapshot of machine learning literature. This study is not a subject specific literature review. We grouped 30 publications are classified into six different topic (clustering, classification, regression, multivariate data analysis and image processing) according to *Table 1*.

We searched Science Driect digital databases to get frequency of machine learning algortihms and publications rate per individual year. We selected keywords: (Veterinary

OR animal) AND (machine learning algorthms which is related in this study). The percentage of publications according to machine learning methods and publications frequency rate per year (2002 - 2016) is shown in *Fig. 4*. The most frequently used methods are NN, LogR, LinR, multiple regression, PCA and k-means. Also we have identified that the literature on machine learning application in veterinary field up and upward momentum with the proliferation.

CONCLUSION

Machine learning is a field of study associated with artificial intelligence, in order to enable a machine to perform a task on its own and with the best performance and this is a discipline based on extracting information and learning from data. A learning computer is able to generate predictions about future thanks to machine learning methods. Therefore, it contributes a great deal to scientific researches by being widely used especially in the field of medicine.

Today, precautions to avoid diseases on animal health and particularly in herd health and herd management in farms, determination of risk factors and development of appropriate protection-control programs are being assessed to be quite important and popular. Multi-disciplinary studies carry a significant potential for contribution to this field of research.

In this paper, related works concerning machine learning applications from veterinary field are reviewed. For this purpose 30 publications are classified into clustering, classification, regression, multivariate data analysis and image processing. The main goal of this review is to introduce a baseline and an idea to researchers who would work on this field. Thus, it is envisaged that the machine learning applications on veterinary shall increase and the latest developments on machine learning shall be applied to solving the problems on animal health.

We have identified that the literature on machine learning application in veterinary field up and upward momentum with the proliferation. And the neural network, logistic regression, linear regression, multiple regression, principle component analysis and k-means methods are most frequently used.

It was observed that recent developments in computer science/engineering have not adequately addressed the problems encountered in veterinary field. Such as deep learning, ensemble learning, voice recognition, emotion recognition, etc. is still new in the field of veterinary.

In this study, we focused machine learning applications in veterinary field. This study is not a subject specific literature review. Therefore as for future works, provide a review on computer aided disease diagnosis in veterinary. Also, introducing the tools, package and successful implemented programs in this field is proposed.

REFERENCES

1. Witten IH, Frank E: Data Mining: Practical Machine Learning Tools and Techniques. 2nd edn., Morgan Kaufmann Publishers, 29-30, 2005.

2. Amasyalı MF: Yeni makine öğrenmesi metodları ve ilaç tasarımına uygulamaları. *Doktora Tezi,* Yıldız Teknik Üniv. Fen Bil. Enst., 2008.

3. Bal M, Sever H, Kalıpsız O: Modeling the symptom-disease relationship by using rough set theory and formal concept analysis. *World Acad Sci Eng Technol (WASET)*, 26 (12): 517-521, 2007.

4. Aktaş MS, Kalıpsız O: Veri madenciliğinde özellik seçim tekniklerinin bankacılık verisine uygulanması üzerine araştırma ve karşılaştırmalı uygulama. *9. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS),* 09-11 Eylül, İzmir, 72-83, 2015.

5. Kırtay SH, Ekmekci N, Halıcı T, Ketenci U, Aktas MS, Kalıpsız O: Pazar sepeti analizi için örneklem oluşturulması ve birliktelik kurallarının çıkartılması. *9. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS),* 09-11 Eylül, İzmir, 172-183, 2015.

6. Cihan P, Kalıpsız O: Öğrenci proje anketlerini sınıflandırmada en başarılı algoritmanın belirlenmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi (TBV BBMD),* 8 (1): 41-49, 2015.

7. Cihan P, Kalıpsız O, Cingiz MÖ, Doksöz M: Yazılım geliştirme dersleri öğrenci projelerinin birliktelik kuralı ile değerlendirilmesi, 7. *Ulusal Yazılım Mühendisliği Sempozyumu (UYMS)*, 25-28 Eylül, İzmir, 2013.

8. Olcaysoy AB, Kalıpsız O, Gürlesin T, Türkdoğan Ş: Yazılım proje faktörlerinin risklerle etkileşimi: Telekomünikasyon örneği. *9. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS),* 09-11 Eylül, İzmir, 47-58, 2015.

9. Ayyıldız M, Kalıpsız O, Sırma Y: Yazılım geliştirme projelerinde yapay sinir ağı kullanarak maliyet tahmini. *III. Ulusal Yazılım Mühendisliği Sempozyumu ve Sergisi (UYMS),* 27-30 Eylül, Ankara, 2007.

10. Hong L, Wan Y, Jain A: Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Trans Pattern Anal Mach Intell*, 20 (8): 777-789, 1998.

11. Liam LW, Chekima A, Fan LC, Dargham JA: Iris recognition using self-organizing neural network. *Res Develop Student Conf IEEE*, 2002, SCOReD 2002,169-172, 2002. DOI: 10.1109/SCORED.2002.1033084

12. Viola P, Jones MJ: Robust real-time face detection. *Int J Comput Vision*, 57, 137-154, 2004. DOI: 10.1023/B:VISI.0000013087.49260.fb

13. Marti UV, Bunke H: Using a statistical language model to improve the performance of an HMM-based cursive handwriting recognition system. *Int J Pattern Recogn*, 15, 65-90, 2001. DOI: 10.1142/S0218001401000848

14. Hoi SCH, Jin R, Zhu J, Lyu MR: Batch mode active learning and its application to medical image classification. *Proceedings of the 23rd International Conference on Machine learning (ICML '06)*, June 25-29, Pennsylvania, USA, 417-424, 2006. DOI: 10.1145/1143844.1143897

15. Cordón O, Herrera-Viedma E, Lapez-Pujalte C, Luque M, Zarco C: A review on the application of evolutionary computation to information retrieval. *Int J Approx Reason*, 34, 241-264, 2003. DOI: 10.1016/j.ijar.2003.07.010

16. Kaur H, Wasan S: Empirical study on applications of data mining techniques in healthcare. *J Comput Sci*, 2 (2): 194-200, 2006.

17. Şentürk ZH: Veri madenciliğiyle kanser tanısı. *Yüksek Lisans Tezi,* Düzce Üniv. Fen Bil. Enst., 2011.

18. Patil MB, Joshi RC, Toshniwal D, Biradar S: A new approach: Role of data mining in prediction of survival of burn patients. *J Med Syst*, 35, 1531-1542, 2011. DOI: 10.1007/s10916-010-9430-2

19. Elmas F: Kalp krizi riskinin bir veri madenciliği uygulaması ile analizi. *Yüksek Lisans Tezi,* Muğla Sıtkı Koçman Üniv. Fen Bil. Enst., 2014.

20. Şık MŞ: Veri madenciliği ve kanser erken teşhisinde kullanımı. *Yüksek Lisans Tezi,* İnönü Üniv. Sosyal Bil. Enst., 2014.

21. Erkuş S: Veri madenciliği yöntemleri ile kardiyovasküler hastalık tahminin yapılması. *Yüksek Lisans Tezi,* Bahçeşehir Üniv. Fen Bil. Enst., 2015.

22. Topuz MD: Makine öğrenmesi algoritmaları ve anomali tespiti. *Yüksek Lisans Tezi,* Bahçeşehir Üniv. Fen Bil. Enst., 2014.

23. Gürcan S, Akçapınar H: Alman et ve karacabey merinosu koyunlarının canlı ağırlık, vücut ölçüleri ve yapağı inceliği yönünden kümeleme analizi ile incelenmesi. *Turk J Vet Anim Sci*, 26, 1255-1261, 2002.

24. Küçükönder H, Üçkardeş F, Narinç D: Hayvancılık alanında bir veri madenciliği uygulaması: Japon bıldırcını yumurtalarında döllülüğe etki eden bazı faktörlerin belirlenmesi. *Kafkas Univ Vet Fak Derg*, 20, 903-908, 2014. DOI: 10.9775/kvfd.2014.11353

25. Ghotoorlar SM, Ghamsari SM, Nowrouzian I, Ghotoorlar SM, Ghidary SS: Lameness scoring system for dairy cows using force plates and artificial intelligence. *Vet Rec*, 170, 126, 2012. DOI: 10.1136/vr.100429

26. Gökçe E, Atakişi O, Kırmızgül AH, Erdoğan HM: Risk factors associated with passive immunity, health, birth weight and growth performance in lambs: III. The relationship among passive immunity, birth weight gender, birth type, parity, dam's health and lambing season. *Kafkas Univ Vet Fak Derg*, 19, 741-747, 2013. DOI: 10.9775/kvfd.2013.8441

27. Bozkurt Y, Aydoğan T, Tüzün CG: Açıkta besi (feedlot) sisteminde yetiştirilen esmer ve siyah alaca ırkı hayvanların sayısal görüntü işleme ve yapay sinir ağları yöntemi ile performans ve karkas özelliklerinin saptanması, Tübitak Projesi, Proje no: 110O269, 2013.

28. Alpaydın E: Introduction to Machine Learning. MIT Press, 2nd ed., 3-4, Cambridge, MA, 2010.

29. Nilsson N: Introduction to Machine Learning. California, United Stated of Americas, 1996.

30. Kudyaba S, Hoptroff R: Data Mining and Business Inteligence. Idea Group Publishing, 2001.

31. Gökçe E, Atakişi O, Kırmızıgül AH, Ünver A, Erdoğan HM: Passive immunity in lambs: Serum lactoferrin concentrations as a predictor of IgG concentration and its relation to health status from birth to 12 weeks of life. *Small Ruminant Res*, 116, 219-228, 2014. DOI: 10.1016/j.smallrumres.2013.11.006

32. Tirozzi B, Bianchi D, Ferraro E: Introduction to computational neurobiology and clustering. 104-106, World Scientific, Singapore, 2007.

33. Schalkoff RJ: Digital Image Processing and Computer Vision. John Wiley, New York, 1989.

34. Hair JR, Joseph F, Hult GTM, Ringle C, Sarstedt M: A primer on partial least squares structural equation modeling (PLS-SEM). Sage Publications, 2016.

35. Akıllı A, Atıl H, Takma Ç, Ayyılmaz T: Fuzzy logic-based decision support system for dairy cattle. *Kafkas Univ Vet Fak Derg*, 22, 13-19, 2016. DOI: 10.9775/kvfd.2015.13516

36. Hempstalk K, McParland S, Berry DP: Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *J Dairy Sci,* 98, 5262-5273, 2015. DOI: 10.3168/jds.2014-8984

37. Lewis FI, Brülisauer F, Gunn GJ: Structure discovery in Bayesian networks: An analytical tool for analysing complex animal health data. *Prev Vet Med*, 100, 109-115, 2011. DOI: 10.1016/j.prevetmed.2011.02.003

38. Karabağ K, Alkan S, Mendeş M: Classification tree method for determining factors that affecting hatchability in chukar partridge (alectoris chukar) eggs. *Kafkas Univ Vet Fak Derg,* 16, 723-727, 2010. DOI: 10.9775/kvfd.2009.1539

39. Ortiz-Pelaez Á, Pfeiffe DU: Use of data mining techniques to investigate disease risk classification as a proxy for compromised biosecurity of cattle herds in Wales. *BMC Vet Res,* 4, 1-16, 2008. DOI: 10.1186/1746-6148-4-24

40. Gokce E, Kırmızıgul AH, Citil M, Erdogan HM: Risk factors associated with passive immunity, health, birth weight and growth performance in lambs: I. Effect of parity, dam's health, birth weight, gender, type of birth and lambing season on morbidity and mortality. *Kafkas Univ Vet Fak Derg,* 19 (Suppl-A): A153-A160, 2013. DOI: 10.9775/kvfd.2012.8440

41. Gökce E, Atakisi O, Kırmızigul AH, Erdogan HM: Risk factors associated with passive immunity, health, birth weight and growth performance in lambs: II. Effects of passive immune status and some risk factors on growth performance during the first 12 weeks of life. *Kafkas Univ Vet Fak Derg*, 19, 619-627, 2013. DOI: 10.9775/kvfd.2013.8442

42. Teke EÇ, Orhan H, Küçüksille EU, Bilginturan S, Teke H: Veri madenciliği süreci ile siyah alaca sığırlarda canlı ağırlık tahmini. *8. Ulusal Zootekni Bilim Kongresi,* 5-7 Eylül, Çanakkale, 365, 2013.

43. Takma Ç, Atıl H, Aksakal V: Çoklu doğrusal regresyon ve yapay sinir ağı modellerinin laktasyon süt verimlerine uyum yeteneklerinin karşılaştırılması. *Kafkas Univ Vet Fak Derg*, 18, 941-944, 2012. DOI: 10.9775/kvfd.2012.6764

44. Hermann-Bank M L, Skovgaard K, Stockmarr A, Strube ML, Larsen N, Kongsted H, Ingerslev HC, Mølbak L, Boye M: Characterization of the bacterial gut microbiota of piglets suffering from new neonatal porcine diarrhoea. *BMC Vet Res*, 11, 1-19, 2015. DOI: 10.1186/s12917-015-0419-4

45. Nantima N, Ocaido M, Ouma E, Davies J, Dione M, Okoth E, Mugisha A, Bishop R: Risk factors associated with occurrence of African swine fever outbreaks in smallholder pig farms in four districts along the Uganda-Kenya border. *Trop Anim Health Prod*, 47, 589-595, 2015. DOI: 10.1007/s11250-015-0768-9

46. Dupuy C, Morignat E, Maugey X, Vinard JL, Hendrikx P, Ducrot C, Calavas D, Gay E: Defining syndromes using cattle meat inspection data for syndromic surveillance purposes: A statistical approach with the 2005-2010 data from ten French slaughterhouses. *BMC Vet Res*, 9, 1-17, 2013. DOI: 10.1186/1746-6148-9-88

47. Warns-Petit E, Morignat E, Artois M, Calavas D: Unsupervised clustering of wildlife necropsy data for syndromic surveillance. *BMC Vet Res,* 6, 1-11, 2010. DOI: 10.1186/1746-6148-6-56

48. Doğan İ: Selection by cluster analysis. *Turk J Vet Anim Sci,* 26 (1): 47-53, 2002.

49. Yunusa AJ, Salako AE, Oladejo OA: Principal component analysis of the morphostructure of Uda and Balami sheep of Nigeria. *Int Res J Agric Sci,* 1 (3): 45-51, 2013.

50. Akçay A, Uğurlu M, Yakan A, Atasoy F: Etçi piliçlerde karkas özelliklerinin temel bileşenler analizi ile değerlendirilmesi. *Uluslararası Katılımlı XIV. Ulusal Biyoistatistik Kongresi,* 91, 2012.

51. Casanova PMP, Sinfreu I, Villalba D: Principal component analysis of cephalic morphology to classify some Pyrenean cattle. *Anim Genet Resour*, 50, 59-64, 2012. DOI: 10.1017/S2078633611000385

52. Meyer K: Multivariate analyses of carcass traits for Angus cattle fitting reduced rank and factor analytic models. *J Anim Breed Genet*, 124, 50-64, 2007. DOI: 10.1111/j.1439-0388.2007.00637.x

53. Lee CH, Chou CH, Han CC, Huang RZ: Automatic recognition of animal vocalizations using averaged MFCC and linear discriminant analysis. *Pattern Recognition Letters*, 27, 93-101, 2006. DOI: 10.1016/j.patrec.2005.07.004

54. Caraviello DZ, Weigel KA, Craven M, Gianola D, Cook NB, Nordlund KV, Fricke PM, Wiltbank MC: Analysis of reproductive performance of lactating cows on large dairy farms using machine learning algorithms. *J Dairy Sci*, 89, 4703-4722, 2006. DOI: 10.3168/jds.S0022-0302(06)72521-8

55. Bilgin ÖC, Esenbuğa N: Morkaraman koyunlarının vücut ölçüleri ve karkas ağırlıkları arasındaki ilişkinin değerlendirilmesinde kanonik korelasyon analizinin kullanılması. *GAP IV. Tarım Kongresi*, 21-23 Ekim, Şanlıurfa, 21-23, 2005.

56. Mcevoy FJ, Amigo JM: Using machine learning to classify image features from canine pelvic radiographs: Evaluation of partial least squares discriminant analysis and artificial neural network models. *Vet Radiol Ultrasound,* 54, 122-126, 2013. DOI: 10.1111/vru.12003

57. Bilgin E, Ceylan M, Yalçın H: A digital image processing based bio-identification application from planum nasale of Kangal dogs. *IEEE 19th Signal Processing and Communications Applications Conference (SIU)*, 20-22 April, Antalya, 275-278, 2011. DOI: 10.1109/SIU.2011.5929640

58. Slósarz P, Stanisz M, Boniecki P, Przybylak A, Lisiak D, Ludwiczak A: Artificial neural network analysis of ultrasound image for the estimation of intramuscular fat content in lamb muscle. *Afr J Biotechnol*, 10, 11792-11796, 2011.