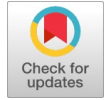# Eye Disease Prediction Among Corporate Employees using Machine Learning Techniques

A. Tamilarasi, T. Jawahar Karthick, R. Dharani, S. Jeevitha

Abstract: In the IT sector, employees use systems for more than 6 hs, so they are affected by many health problems. Mostly In the IT sector, employees are affected with eye diseases like eye strain, eye pain, burning sensation, double vision, blurring of vision, and frequent watering. The major goal of this research is to identify the different types of eye problems encountered, the symptoms present, and the population afflicted by eye diseases in order to accurately forecast outcomes using a Machine learning techniques for real-time data sets.

Keywords: Machine Learning Techniques

I. INTRODUCTION

The eye is the smallest and one of the most important organs in the human body. Therefore, it is crucial to take care of it. As a result, since the majority of diseases have a brain component, it is essential to anticipate ocular issues, which calls for comparative research. Since the inaccuracy of the instrument causes the majority of patients to lose their eyes nowadays, it is crucial to comprehend the most efficient ways to reduce illness risk. The testing approach that has had the most success is machine learning. In the vast field of study known as machine learning (ML), computers are taught to mimic human talents. The term "machine intelligence," which describes the fusion of the two technologies, refers to machine learning systems that are taught how to analysis and utilize data. As testing data for this paper, use bio factors like

- Age
- Experience
- H of work

evaluate the accuracy with the use of the following algorithms, Naive Bayes, SVM, KNN. In this study, predict the accuracy of above three different machine learning algorithms and compare which one is the best.

Dr. A. Tamilarasi, Professor, Department of Computer Applications, Kongu Engineering College, Perundurai-638060, (Tamilnadu), India. E-mail: drtamil@kongu.ac.in, ORCID ID: 0000-0002-5885-5541

T. Jawahar Karthick, PG Final Year, Department of Computer Applications, Kongu Engineering College, Perundurai-638060, (Tamilnadu), India.Email: jawaharkarthickt.21mca@kongu.edu

R. Dharani, PG Final Year, Department of Computer Applications, Kongu Engineering College, Perundurai-638060, (Tamilnadu), India. E-mail: dharanir.21mca@kongu.edu

S. Jeevitha, PG Final Year, Department of Computer Applications, Kongu Engineering College, Perundurai-638060, (Tamilnadu), India. E-mail: jeevithas.21mca@kongu.edu

The two ideas of testing and training are the foundation of the potent instrument known as machine learning. The application of a test to various types of needs is done by a system using the information it has gained through data and experience and an algorithm. Supervised, unsupervised, and reinforced machine learning approaches are the three categories.

II. SUPERVISED LEARNING

Supervised Learning is defined as learning with the assistance of a teacher or a qualified guide. There is always a training dataset available when testing data is available since have a database that can be used to teach prediction in a specific database. "Train me" is the guiding premise of supervised learning. The following are the fundamental steps of supervised learning:

- Classification
- Regression

Regression is a method for identifying patterns and estimating the likelihood of outcomes that cannot be changed. a system for classifying numbers' senses, values, and other numbers like width and height.

III. UNSUPERVISED LEARNING

Unsupervised Learning is defined as learning that takes place while an instructor is not there to provide guidance. Unsupervised learning automatically analyses a dataset to find patterns and connections between them. When fresh data is added, it classifies it and stores it in one of the connections. Unsupervised learning is motivated by the idea of "self-sufficiency".

Imagine you have a variety of fruits, such as mango, banana, and apple, and we use unsupervised learning to group them into three distinct clusters based on how they are related to one another. When new data is received, it is automatically sent to one of the clusters.

IV. REINFORCEMENT

The ability of an agent to interact with their surroundings and choose the desired outcome is referred to as reinforcement. The "hit and trial" idea underlies reinforcement. In reinforced learning, each agent receives positive and negative points, with positive points being given first.

V. EYE PROBLEM

Most people experience eye issues at some point. Some are simple to cure at home or are minor and will go away on their own.

Others need the assistance of an expert. There are things we can do to improve eye heal been regardless of whether we currently have bad vision or never have. See whether any of these common problems ring a bell with us

A. Red Eyes:

Blood vessels that cover their surface enlarge when they get inflamed or diseased. Our eyes appear red as a result of that.Allergies, eyestrain, late nights, little sleep, or allergies may be the cause. Consult a doctor if an injury is the root of the problem. Red eyes may be a sign of another ocular disorder, including conjunctivitis (conjunctivitis) or UV damage caused by years of not using sun glasses. Consult the doctor if non-prescription eye drops and rest do not alleviate y symptoms.

B. Eyestrain:

This one is ll known to anyone who spends a lot of time reading, using a computer, or traveling long distances. When we use our eyes excessively, it occurs. Like any other part of body, they grow fatigued and need to rest. Give our eyes a break if they are feeling tired. After few days, if they are still tired, speak to y doctor to be sure that there is no other problem.

C. Night Blindness:

Night blindness is what that sounds like. It is merely a symptom and not a distinct issue. One kind of night blindness that will be treated by doctors is caused by myopia, opaque.

D. Lazy Eye:

This problem can be inherited, and it can also be brought on by a retinal degenerative disorder that is often incurable. If we have lazy eye we must exercise particular caution in dimly lit areas.Amblyopia, often known as "lazy eye," develops when one eye is not correctly formed. That eye has poorer vision and moves more "lazily" than the other while remaining still. Infants, kids, and adults can develop it, but seldom do both eyeInfants and youngsters need to get treatment right away. If it is identified and handled in starting stage, lifelong visual issues can be prevented.

VI. NYSTAGMUS

One might have strabismus if, when they gaze at something, their eyes aren't aligned with one another. Crossed eyes and walleye are some names for it.This issue won't just go away by itself. To help reinforce eye muscle akness, may attend eye therapy sessions with an ophthalmologist from time to time.The surgical treatment will typically need to be done by an ophthalmologist or eye surgeon expert.

A. Uveitis:

This is the term used to describe a group of conditions that inflame the uvea. The majority of the blood vessels are located in the central layer of the eye. These conditions can damage ocular tissue and potentially result in the loss of an eye. It is available to all ages of people. Symptoms could disappear right away or might linger for a while. People suffering from diseases of the immune system.

VII. PRESBYOPIA

This occurs when our eyesight at a distance is adequate but we have trouble seeing small print and close-up things effectively. From the age of 40, the need to hold books and learning materials at a They may be read more easily from a distance.T o regain clear reading vision, one can have LASIK laser eye surgery, contact lenses, reading glasses, and other procedures.

A. Floaters:

They glide into y area of vision as tiny specks or dots. Most people become aware of them when they are outside on a sunny day or in ll-lit settings. Floaters are typically normal, but they can also indicate a more serious eye condition, like a detached retina.

B. Dry Eyes:

When our eyes are unable to produce enough high-quality tears, this occurs. Our eyes may feel as though they are burning or that something is in them. Extreme dryness can occasionally, in severe situations, cause some visual loss. Some remedies consist of:

- Use a home humidifier
- Unique eye drops that mimic the effects of actual tears
- Tear duct plugs that reduce drainage
- Dry eye treatment called Lipiflow involves heat and pressure

might have dry eye disease if frequently have dry eyes. To increase tear production, doctor may advise using cyclosporine (Restsis) or Xiidra eye drops.

C. Excess Tearing:

Tears can be a sign of more serious medical conditions, such as, eye infection or tear duct blockage. An eye doctor can treat or cure these problems

D. Cataracts:

In the eye lens, these manifest as blurry spots. Like a camera lens, a healthy lens is transparent. It allows light to reach our retina, which is in the rear of our eye and is responsible for processing images. Cataracts impair the transmission of light.

E. Glaucoma

Like a tire, our eye has some safe and typical pressure inside of it. Hover, excessive doses can harm our visual nerve. The term "glaucoma" refers to a set of illnesses that induce this condition. Primary open angle glaucoma is a typical kind. Most patients don't experience any early signs of illness or suffering. Therefore, it's crucial to maintain y monthly eye checkups. Even though it's uncommon, glaucoma can be brought on by:

- Anhurts the eyes
- Vascular blockages

F. Conjunctivitis (Pinkeye):

This condition causes inflammation of the tissue that lines the sclera and the rear of our eyelids.

It might result in redness, tears, discharge, itching, burning, or the sensation that something is in our eye. All ages can purchase it. Infections, contact with irritants and chemicals, or allergies are some of the causes. Wash our hands frequently to reduce the risk of contracting it.

VIII. RELATED WORKS:

In this section, we discuss about the papers related to prediction of ocular problems using AI and ML using five eye conditions described below. Some of the results from these studies are utilized for a more accurate interpretation of the findings from the proposed approach, numerous papers outlining comprehensible AI techniques are also provided. also review small studies conducted in the field of eye disease analytics based on symptom dataset.

A study looked at the prevalence of eye disease among low-wage people in Bangladesh. The study used in-person questionnaires, and the results showed that eye diseases are more common than the norm worldwide. This represents a discrepancy in the incidence and diagnosis of ocular disease among low-wage and high-living individuals [1]. The paper claims that there is a communication and financial divide beten the Bangladeshi community as a whole and Bangladeshi slum residents who require aid with their vision. The Shahjadpur Children's Cohort (SCC), a particularly fascinating population-based cohort composed up entirely of children, is the subject of a study on the epidemiology of ocular diseases in reference [2]. discovered some articles outlining approaches to classifying eye diseases using ML algorithms and image processing methods. Train the version the use of photograph processing capabilities which can pick out areas of interest (ROI) [3]. in the image, these techniques require a sizeable image data set

for classification. Sakri et al. [4] an automatic classification framework for diagnosing diabetic eye disease using image processing in one study (DED). Lighting adjustments, CLAHE, and image separation are all methods for enhancing

Table 1: Features of Dataset

S.no	Attribute	Descriptions	Type
1	Age	Age of the Employee	Numerical data
2	Sex	Employee Gender	Nominal data
3	Experience	No of year Experience in employee	Numerical data
4	Spend Hs	Employees working hs	Numerical data
5	Eye Strain	Classification of symptoms	Numerical data
6	Burning Sensation	Types of eye problem	Numerical data

photographs. Macular region, optic nerve, and vessel detection methods re used before pre-trainmodels (VGG-16, Exception, ResNet50, CNN) re used. All models performed admirably with over 90% accuracy. A similar article on image processing with different datasets can be found in [5–8].A relatively typical case in the detection of eye illnesses uses a neural network-based technique using an image dataset. On the Aptos-2019 and IDrID datasets, Nazir et alimproved.'s CenterNet approach was used to extract features using

DenseNet-100. The Aptos-2019 dataset and IDrID dataset both contributed to the method's greatest accuracy results, which re 97.93% and 98.10%, respectively. [9] applied feature fusion techniques and deep neural networks on the same Aptos 2019 dataset. The accuracy rate for this method is 84.31%. Khan and colleagues [10] sought to manually extract the retinal features without utilizing feature-selection techniques. In this research, CNN and VGG-19 re combined, and accuracy of 97.47% was attained. CNN was also used by Sarki et al. [11] and Pahuja et al. [12] for picture datasets, and in both cases, accuracy was less than 90%. Malik et al. [13] adopted a data-driven methodology for classifying eye diseases.

A. Data Collection:

Collecting data and choosing a training and testing dataset are the first steps in creating a prediction system. The real-time dataset is collected with a total of 14 properties

B. Data Analysis:

Data analysis is the process of dissecting, changing, reporting, and modeling data to uncover pertinent information, aid in decision-making, and present findings. The process of displaying data in the form of maps or charts is known as data visualization. This makes it considerably simpler to understand large-scale, complex data. It enables decision-makers to construct decisions with greater expertise and to more quickly recognize new trends and patterns. In high-level data analysis for machine learning and research data analysis, it is also employed (EDA). Various tools, such as Python, can be used for data visualization. In this paper we used Py Charm package to conduct data analysis. All of the target properties will be examined.

IX. ATTRIBUTE SELECTION

The characteristics of datasets are those of the datasets utilized by computers, and different attributes of the eye, including a person's experience, sex, age, and other factors, are shown in Table 1 for the prediction systems:

A. Train-Test-Split

The train-test split process is used to measure the performance of the build prediction. This is a fast and easy procedure, the solution which allows us to analogize the accomplishment of machine learning algorithms for the anticipative modeling complication. The process has an important structure parameter, which is the size of the train and test packages. It's usually shown as a percentage between 0 and 1 in a train or test datasets:
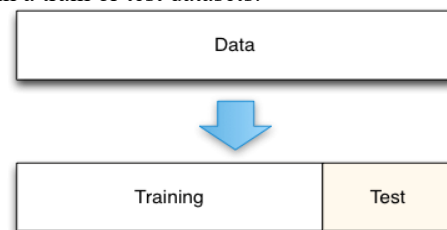
Fig.1 Split up of Data Set

B.   Naive Bayes:

The Naive Bayes technique, a supervised learning algorithm, applies the Bayes theorem to taxonomy problems. It has a sizable training database and is primarily used for text classification. Naive Bayes is used to forecast the likelihood of various classes based on various attributes. The Bayes theorem provides a principled way to calculate conditional probability. The simplest form of calculation for the Bayes theorem is as follows:

- To import Gaussian NB from the sklearn. naive_bayes package to implement the Naive Bayes algorithm.
- We need to initialize GaussianNB() to
  variable nbnb = GaussianNB()
- We must use fit () to adjust the ight of data values that provide the better accuracy and pass the X_train and Y_train values as a parameter in the function nb.fit(X_train,Y_train)
- To predict data values, you must use Predict() and pass X_test values in the function
  Y_pred_nb =
  nbnb.predict(X_test)

To calculate an accurate score, you must use accuracy_score (), pass the predicted data values and _test values as parameters into the function, and store the value of accuracy_score in the variable score_nb

- At the end of this process you will get the accuracy value of the Naive based algorithm

Output:
The accuracy score achieved using Naive Bayes is: 51.92 %

SVM:

The most common Supervised Learning technique for handling classifying and regression problems is the Support Vector Machine. In machine learning, it is typically used to address classification issues.

- We need to import SVM from the sklearn package to implement the SVM algorithm.
- Then   initialize   SVM.SV   =
  svm.SVC(kernel='linear') SVC() to variable SV.
- Use fit () to adjust the weight of data values that provide the better accuracy and pass the X_train and Y_train values as a parameter in the function
  SV.fit(X_train,Y_train)
- To predict data values, you must use Predict() and pass X_test values as parameter in the function
  Y_pred_svm = sv.predict(X_test)

To calculate an accurate score, we must use accuracy_score(), pass the predicted data values.

- Y_test values as parameters into the function, and store the value of accuracy_score in the variable score_svm
- At the end of this process we will get the accuracy value of the SVM algorithm

Output:
The accuracy score achieved using Linear SVM is: 70.13 %
KNN:
The KNN is one of the crucial basic jobs in machine learning. KNN is based on the Supervised Learning method. The KNN technique, which assumes that fresh data and earlier cases are similar, places new cases in categories that best fit older categories. The KNN algorithm gathers all of the data that is accessible and organizes fresh data points according to similarity. As soon as new data is introduced, the K-NN

algorithm can instantly classify pertinent packet kinds. Large amounts of training data might make it more efficient.

- We need to import K-Neighbors Classifier from the sklearn.
- Neighbors package to implement the KNN algorithm.
- Initialize
   K-Neighbors Classifier (n_neighbors=7) to variable knn
  Knns= K-Neighbors Classifier()
- We must use fit () to adjust the weight of data values that provide better accuracy and pass the X_train and Y_train values as a parameter in the function
  knns.fit(X_train,Y_train)
- To predict data values, we must use Predict() and pass X_test values as parameter in the function
  Y_pred_knn=knns.predict(X_test)
- To calculate an accurate score, we must use accuracy_score (), pass the predicted data values and y_test values as parameters into the function, and store the value of accuracy_score in the variable score_knn
- At the end of this process, we will get the accuracy value of the KNN algorithm.

Output:
The accuracy score achieved using KNN is: 89.25 %

Table 2: Comparison of Algorithms

Algorithm	Accuracy
KNN	89.25%
SVM	70.13%
NAÏVE BAYES	51.92%

X.   CONCLUSION

Due to the importance of the eye as a crucial organ and the fact that predicting eye diseases is a major worry for people, algorithm accuracy is one of the factors used to evaluate algorithm performance. The dataset used for training and testing machine learning algorithms affects how accurate they are. KNN gives better accuracy compare to other algorithms, for the given dataset with the properties mentioned inTable 2

DECLARATION STATEMENT

Funding/ Grants/ Financial Support	No, I did not receive.
Conflicts of Interest/ Competing Interests	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval and consent to participate with evidence.
Availability of Data and Material/ Data Access Statement	Not relevant.
Authors Contributions	I am only the sole author of the article.

4

REFERENCES

1. An, G.; Omodaka, K.; Tsuda, S.; Shiga, Y.; Takada, N.; Kikawa, T.; Nakazawa, T.; Yokota, H.; Akiba, M. Comparison of machine-learning classification models for glaucoma management. J. Health. Eng. 2018, 2018. https://doi.org/10.1155/2018/6874765
2. Agrawal, P.; Madaan, V.; Kumar, V. Fuzzy rule-based medical expert system to identify the disorders of eyes, ENT and liver. Int. J. Adv. Intell. Paradig. 2015, 7, 352–367. https://doi.org/10.1504/IJAIP.2015.073714
3. Sample, P.A.; Boden, C.; Zhang, Z.; Pascual, J.; Lee, T.W.; Zangwill, L.M.; inreb, R.N.; Crowston, J.G.; Hoffmann, E.M.; Medeiros, F.A.; et al. Unsupervised machine learning with independent component analysis to identify areas of progression in glaucomatous visual fields. Investig. Ophthalmol. Vis. Sci. 2005, 46, 3684–3692. https://doi.org/10.1167/iovs.04-1168
4. Imberman, S.P.; Ludwig, I.; Zelikovitz, S. Using Decision Trees to Find Patterns in an Ophthalmology Dataset. In Proceedings of the FLAIRS Conference, Palm Beach, FL, USA, 18–20 May 2011.
5. Arbelaez, M.C.; Versaci, F.; Vestri, G.; Barboni, P.; Savini, G. Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data. Ophthalmology 2012, 119, 2231–2238. https://doi.org/10.1016/j.ophtha.2012.06.005
6. Fageeri, S.O.; Ahmed, S.M.M.; Almubarak, S.A.; Muazu, A.A. Eye refractive error classification usingmachine learning techniques. In Proceedings of the IEEE International Conference on Communication,Control, Computing and Electronics Engineering, Khartoum, Sudan, 16–17 January 2017; pp. 1–6. https://doi.org/10.1109/ICCCCEE.2017.7867660
7. Organization, W.H. International Classification of Diseases (ICD). Available online: http://www.who.int/ classifications/icd/ICD10Volume2_en_2010.pdf (accessed on 1 January 2017).
8. Waudby, C.J.; Berg, R.L.; Linneman, J.G.; Rasmussen, L.V.; Peissig, P.L.; Chen, L.; McCarty, C.A. Cataract research using electronic health records. BMC Ophthalmol. 2011, 11, 32. https://doi.org/10.1186/1471-2415-11-32
9. Sullivan, B.D.; Crews, L.A.; Messmer, E.M.; Foulks, G.N.; Nichols, K.K.; Beginninger, P.; Geerling, G.; Figueiredo, F.; Lemp, M.A. Correlations beten commonly used objective signs and symptoms for the diagnosis of dry eye disease: Clinical implications. Acta Ophthalmol. 2014, 92, 161–166. https://doi.org/10.1111/aos.12012
10. Moccia, S.; De Momi, E.; El Hadji, S.; Mattos, L.S. Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics. Comput. Methods Programs Biomed. 2018, 158, 71–91. https://doi.org/10.1016/j.cmpb.2018.02.001
11. Fraz, M.M.; Remagnino, P.; Hoppe, A.; Uyyanonvara, B.; Rudnicka, A.R.; On, C.G.; Barman, S.A. Blood vessel segmentation methodologies in retinal images–a survey. Comput. Methods Programs Biomed. 2012, 108, 407–433. https://doi.org/10.1016/j.cmpb.2012.03.009
12. Quellec, G.; Lamard, M.; Erginay, A.; Chabouis, A.; Massin, P.; Cochener, B.; Cazuguel, G. Automatic detection of referral patients due to retinal pathologies through data mining. Med. Image Anal. 2016, 29, 47–64. https://doi.org/10.1016/j.media.2015.12.006
13. Burgansky-Eliash, Z.; Wollstein, G.; Chu, T.; Ramsey, J.D.; Glym, C.; Noecker, R.J.; Ishikawa, H.; Schuman, J.S. Optical coherence tomography machine learning classifiers for glaucoma.

AUTHOR PROFILE

Dr. A. Tamilarasi is currently working as Professor in MCA at Kongu Engineering College, Erode, TamilNadu. She has 27 years of Teaching and 33 years of research experience and has supervised 33 Ph.D. and 6-M.Phil candidates. She has 1 granted patent, 10 books, more than 180 research papers in Scopus/SCI journals of repute, more than 50 national and international conference papers, and 5 awards to her credit. She completed research projects sponsored by UGC and ICMR and organized many sponsored programmes like FDPs and conferences. She is a reviewer for more than 20 reputed journals and a life member of ISTE.

Mr. Jawahar Krthick is doing his MCA course in Kongu Engineering College. His area of interest includes Machine Learning and IoT. He presented papers in International Conferences and completed two projects. His strength includes hard working, well disciplined and ability to complex solves problems very quickly.

Ms. R. Dharani is doing his MCA course in Kongu Engineering College. Her area of interest includes AI and cloud computing. She presented papers in various Conferences and completed a project. To use her relevant experience, technical knowledge, and problem-solving abilities to land a professional job in an MNC. She is able to speak effectively with a group, and an excellent learner and listener. She inspires friends and also focus mainly to achieve goals. She is so active both in academic and research

S. Jeevitha, as a postgraduate student at Kongu Engineering College, Ms. Jeevitha is passionate about backend programming languages like Python, Java, JavaScript, and PHP. She works hard in order to obtain a professional position at an MNC by utilising her relevant experience, technical expertise, and problem-solving skills. She focuses on achieving objectives. She is very active in research and academia.