EasyChair Preprint

№ 10623

Eye Disease Prediction Among Corporate Employees Using Machine Learning Techniques

C Dharani, A Tamilarasi, K Chitra, T Jawahar Karthick and S Jeevitha

July 27, 2023

Eye Disease Prediction Among Corporate Employees Using Machine Learning Techniques

. R.Dharani[1]
PG Final Year
Department of Computer Applications
Kongu Engineering College
Perundurai-638060, Tamilnadu, India.

Dr.A.Tamilarasi [2]
Professor (Sr.G)
Department of Computer Applications
Kongu Engineering College
Perundurai-638060, Tamilnadu, India.

K.Chitra, [3]
Assistant Professor,
Department of Computer Applications
Kongu Engineering College
Perundurai-638060,Tamilnadu, India.

T.Jawahar Karthick[4]
PG Final Year
Department of Computer Applications
Kongu Engineering College
Perundurai-638060, Tamilnadu, India.

S.Jeevitha[5]
PG Final Year
Department of Computer Applications
Kongu Engineering College
Perundurai-638060, Tamilnadu, India.

Abstract— **In the IT sector, employees use systems for more than 6 hours, 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—Eye Disease;Machine Learning.

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. Machine Learning is an area of artificial intelligence. 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 project, we use bio factors like:

- Age
- Experience
- Hour of work

We evaluate the accuracy with the use of the following algorithms:

- Naive Bayes
- SVM
- KNN

In this study, predict the accuracy of three different Machine Learning algorithms and use the results to decide which one is the best. 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.

A. 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 we 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:

- Random Forest
- Classification
- Regression
- Decision Tree

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.

B. Unsupervised Learning

Unsupervised Learning is defined as learning that takes place while an instructor is not there to Unsupervised Learning automatically analyses a dataset to find patterns and connections between them. provide guidance. 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 you 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.

C. 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 Reinforcement Learning, each agent receives positive and negative points, with positive points being given first.

II. 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 you can do to improve your eye heal been regardless of whether you currently have bad vision or never have. See whether any of these common problems ring a bell with you.

A. Red Eye

Blood vessels that cover their surface enlarge when they get inflamed or diseased. Your 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. Redeyes may be a sign of another ocular disorder, including conjunctivitis (conjunctivitis) or UV damage caused by years of not using sunglasses.Consult your doctor if non-prescription eye drops and rest do not alleviate your symptoms.

B. Eye Strain

This one is well known to anyone who spends a lot of time reading, using a computer, or traveling long distances. When you use your eyes excessively, it occurs. Like any other part of your body, they grow fatigued and need to rest.Give your eyes a break if they are feeling tired. After a few days, If they are still tired, speak to your 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 you have it, you 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.

E. Nystagmus

You might have strabismus if, when you gaze at something, your 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 weakness, you 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.

F. 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.

G. Presbyopia

This occurs when your eyesight at a distance is adequate but you 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. As if your arms weren't long enough.To regain clear reading vision, one can have LASIK laser eye surgery, contact lenses, reading glasses, and other procedures.

H. Floaters

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

I. Dry Eye

When your eyes are unable to produce enough high - quality tears, this occurs. Your 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

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

J. Excess Tearing

Your feelings have nothing to do with it. 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.

K. Cataracts

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

L. Glaucoma

Like a tire, your eye has some safe and typical pressure inside of it. However, excessive doses can harm your 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 your monthly eye checkups.
Even though it's uncommon, glaucoma can be brought on by:

- Anhurts the eyes
- Vascular blockages

M. Conjunctivitis (Pink Eye)
This condition causes inflammation of the tissue that lines the sclera and the rear of your eyelids. It might result in redness, tears, discharge, itching, burning, or the sensation that something is in your eye. All ages can purchase it. Infections, contact with irritants and chemicals, or allergies are some of the causes. Wash your hands frequently to reduce your risk of contracting it.

III. LITERATURE REVIEW

The publications related predicting ocular problems using AI and ML are evaluated here because these domains cover the technologies used for the recommended technique to correctly predict the five eye conditions described below. Since 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. We 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 were 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 between 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] We 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).in the image, these techniques require a sizeable image data set for classification. Sakri et al. [3] 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 photographs. Macular region, optic nerve, and vessel detection methods were used before pre – train models (VGG - 16, Exception, Res Net50, CNN) were used. All models performed admirably with over 90% accuracy. A similar article on image processing with different datasets can be found in. [4 – 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 Center Net approach was used to extract features using Dense Net - 100. The Aptos -

2019 dataset and IDrID dataset both contributed to the method's greatest accuracy results, which were 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 were 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.

IV. DATA COLLECTION

Collecting data and choosing a training and testing dataset are the first steps in creating a prediction system. Data collection, for which we use the

Real - time dataset with a total of 14 properties.

A. 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

S.NO	ATTRIBUTE	DESCRIPTION	TYPE
1	Age	Age of the Employee	Numerical Data
2	Sex	Employee Gender	Nominal Data
3	Experience	Employee Experience	Numerical Data
4	Spend Hours	Employee Working Hours	Numerical Data
5	Eye Strain	Classification of Symptoms	Numerical Data
6	Burning Sensation	Types of Eye Problem	Numerical Data

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 essay, we'll use the PyCharm package to conduct data analysis. All of the target properties will be examined.

B. 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:

C. 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 you to analogize the accomplishment of Machine Learning algorithms for your 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:

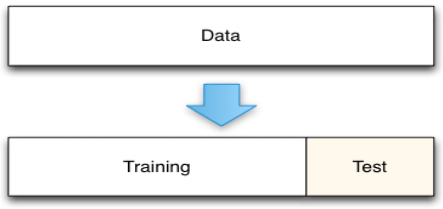

Fig.3 Train Test Spirit

V. ALGORITHMS

A. Naive Bayes

The Naïve 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:

- You need to import Gaussian NB from the sklearn.naive_bayes package to implement the Naive Bayes algorithm.
- You need to initialize Gaussian NB()

  to variable nbnb = Gaussian NB()

- You must 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
  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 Bayes algorithm.

Output

The accuracy score achieved using Naive Bayes is: 51.92 %

B. 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.

- You need to import SVM from the sklearn package to implement the SVM algorithm.
- You need to initialize,
  SVM.SV = svm.SVC(kernel = 'linear')
  SVC() to variable SV.
- You must 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, you 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 you will get the accuracy value of the SVM algorithm.

Output

The accuracy score achieved using Linear SVM is: 70.13 %

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

- You need to import K - Neighbors Classifier from the sklearn.
- Neighbors package to implement the KNN algorithm.
- You need to initialize,
  K - Neighbors Classifier (n_neighbors = 7) to variable knn
  Knns = K – Neighbors Classifier()
- You 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, you 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, you 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, you will get the accuracy value of the KNN algorithm.

Output

The accuracy score achieved using KNN is: 89.25 %

VI. RESULT

ALGORITHM	**ACCURACY**
KNN	89.25%
SVM	70.13%
NAIVE BAYES	51.92%

VII. CONCLUSION

Due to the importance of the eye as a crucial organ and the fact that predicting eye disease 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 is the best algorithm, according to our analysis of algorithms using a dataset with the properties mentioned in TABLE.

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. [CrossRef] [PubMed]

[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. [CrossRef]

[3] Sample, P.A.; Boden, C.; Zhang, Z.; Pascual, J.; Lee, T.W.; Zangwill, L.M.; Weinreb, 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. [CrossRef] [PubMed]

[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. [CrossRef] [PubMed]

[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.

[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. [CrossRef] [PubMed]

[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 between commonly used objective signs and symptoms for the diagnosis of dry eye disease: Clinical implications. Acta Ophthalmol. 2014, 92, 161 – 166. [CrossRef] [PubMed]

[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. [CrossRef] [PubMed]

[11] Fraz, M.M.; Remagnino, P.; Hoppe, A.; Uyyanonvara, B.; Rudnicka, A.R.; Owen, C.G.; Barman, S.A. Blood vessel segmentation methodologies in retinal images–a survey. Computational Methods Programs Biomed. 2012, 108, 407 – 433. [CrossRef]

[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. [CrossRef] [PubMed]

[13] Burgansky - Eliash, Z.; Wollstein, G.; Chu, T.; Ramsey, J.D.; Glymour, C.; Noecker, R.J.; Ishikawa, H.; Schuman, J.S. Optical coherence tomography machine learning classifiers for glaucoma.