



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 6, June 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.542



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Osteoporosis Fracture Analysis Using Machine Learning

Prof. Veena Patil , Madhura Khetre , Priyanka Gayakhe , Diksha Shinde , Mayuri Wankhare

Department Of Computer Engineering, Jayawantrao Sawant College of Engineering, Pune, Maharashtra, India

Computer Engineer, Jayawantrao Sawant College of Engineering, Pune, Baramati, Maharashtra, India

Computer Engineer, Jayawantrao Sawant College of Engineering, Pune , Ahmed Nagar, Maharashtra, India

Computer Engineer, Jayawantrao Sawant College of Engineering, Pune, Beed, Maharashtra, India

Computer Engineer, Jayawantrao Sawant College of Engineering, Pune, Aurangabad, Maharashtra, India

ABSTRACT: Osteoporosis is a bone disease, causing an increased risk of fracture, and is characterised by low bone mineral densities and degradation of bone tissue microarchitecture. Worldwide, one in 3 women and one in 5 men aged more than 50 years are at risk for an osteoporotic fracture. An osteoporotic fracture is predicted to occur every 3 seconds. The most frequent fracture is hip, spine and wrist; osteoporosis develops and is related to it. The likelihood of these fractures rises with greater ageing in both men and women in the hip and spine in particular. Osteoporosis, simply a porous bone, is a condition that reduces bone density and quality. As the bones become more porous and brittle, the chance of fracture increases. Little by little and silently, bone loss begins. There are typically no signs until the initial fracture occurs. The proposed comprehensive approach aims to support osteoporosis therapists to prevent needless bone densitometry testing. In this research, a verified and accurate Machine Learning model is developed to predict the risk of osteoporotic fracture based on a range of variables, including age, sex, health history, and so on.

KEYWORDS: osteopenia, osteoporosis, machine learning, data pre-processing, data mining, evolutionary algorithms, regression, classification, deep learning

I. INTRODUCTION

Osteoporosis is the most prevalent disease of the bone characterised by low bone density mass and structural changes that lower bone tolerance and increase the risk of fracture. During osteoporosis, the density of bone minerals (BMD) disrupted by bone micro-architecture is reduced while the concentration and variety of proteins in bones is reduced. The osteoporotic fractures of the hip, vertebral and wrist are the most frequent. Osteoporotic fractures are characterised as those that have increased often after age 50 at a low BMD location. [\[1\]](#)[\[2\]](#)[\[3\]](#)[\[4\]](#)

II. PROBLEM STATEMENT

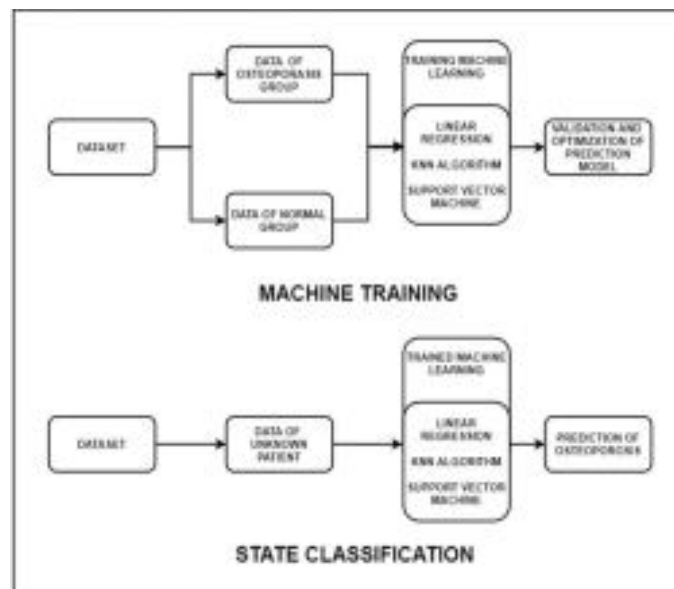
Osteoporosis is defined by an abnormal loss of bone mineral content, which can lead to nontraumatic bone fractures or structural deformations. As a result, bone density has been regarded as the most dependable indicator of osteoporotic fracture risk. Bone texture measurements, by the way, have been investigated in the previous decades to determine other aspects of bone quality. Trabecular structure analysis has been used in several studies to assess bone quality using CT or MR images. To measure trabecular structure, some investigations have used plain x-ray pictures or ultrasound imaging. The majority of studies, however, focus on individual parameters to differentiate between osteoporotic fractured and non fractured groups. Our project will develop an accurate and validated Machine Learning model to predict the risk of osteoporotic fracture based on a dataset that includes age, gender, medical history, and other factors. On the existing dataset, many machine learning models will be implemented in order to determine the most efficient model.

A. PROPOSED METHODOLOGY

1. The task is divided into three main phases: first phase, second stage, last phase.
2. Data Exploration, data cleaning and data transformation are the initial phases.

3. Modeling of data takes centre stage.
4. Data analysis is performed in the last step utilising three models: KNN Algorithm, Linear Regression, and SVM.
5. Instead of using typical data management tools, data exploration is comparable to initial data analysis, visual exploration to learn what is in a dataset and the properties of the data.
6. The process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database is known as data cleaning. It involves identifying incomplete, incorrect, inaccurate, or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data.
7. Data transformation is the process of changing data from one format to another, usually from a source system's format to a destination system's needed format.
8. We move on to data modelling once the first stage is completed. The act of creating a descriptive diagram of links between various types of information to be stored in a database is known as data modelling. One of the goals of data modelling is to come up with the most effective way of storing data while yet allowing full access and reporting.
9. The data is then processed by algorithms, and the results are obtained.
10. These are the test results obtained by using the train dataset to train the models.
11. We can utilise the actual dataset to predict osteoporosis once the dataset has been analysed.

III. SYSTEM ARCHITECTURE



IV. ALGORITHM (MATHEMATICAL MODEL)

4.1 Linear Regression

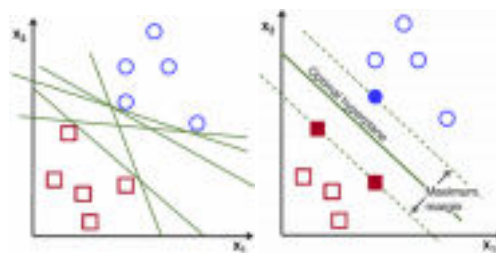
Quarter, upper, normal, and below are all properties in the property rate database. The upper half contains normal estimates for buildings with high expenses, while the lower section has normal estimates for houses in the middle range and low range. The quarter characteristic is plotted on the x-axis, with rate estimates plotted on the y-axis, in order to use straight relapse. A direct relapse is conducted once for each of the quality levels. The decision available to the client to choose from a dropdown list represents the x-axis being autonomous. We accept that there is

a link between the autonomous variable vector and the needy target variable in Linear Regression. We can predict the objective variable by using the free parameters. A vector of N parameters or qualities can be used as the autonomous information vector. They're also known as regressors. It acknowledges a straightforward relationship between the subordinate variable and the regressors. Blunder is a term used to describe the tension between expected and observed regard. The next step is to find the best-fitting relationship (line) between the variables. The Residual Sum of Squares is the most well-known technique (RSS). This method determines the difference between watched information (true esteem) and its vertical separation from the best-fitting line proposed (anticipated esteem). It eliminates all distinctions and encompasses all of them. The MSE (Mean Squared Error) is a quality measure for the estimator that is calculated by separating RSS into watched information foci and adding them up. It is, without a doubt, a non-negative number. Qualities that are closer to 0 indicate a minor error. The square base of the MSE is the RMSE (Root Mean Squared Error). The usual deviation of the assessments from the watched esteems is measured by the RMSE. In comparison to MSE, which might be a large number, this is easier to watch. Unlike other models that can only classify the output, Linear Regression can forecast the precise numerical goal value. As a result, Linear Regression plays a significant role in predicting real estate property market worth.

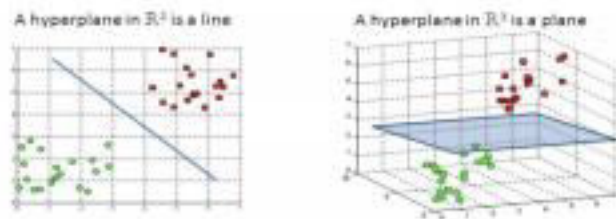
4.2 SVM

The support vector machine algorithm's goal is to find a hyperplane in an N -dimensional space (N — the number of characteristics) that categorises the data points clearly. There are numerous hyperplanes from which to choose to split the two kinds of data points.

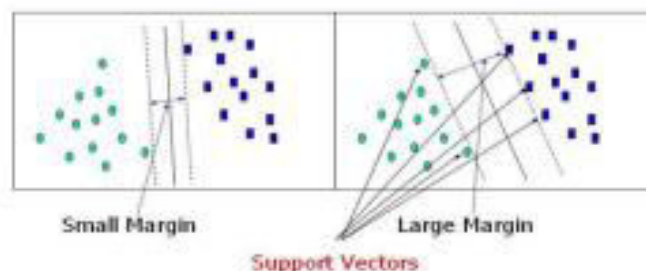
Our goal is to discover a plane with the greatest margin, or the greatest distance between data points from both classes. Maximizing the margin distance gives some reinforcement, making it easier to classify subsequent data points.



hyperplanes and support vectors



Hyperplanes are decision boundaries that aid in data classification. Different classes can be assigned to data dimension is also determined by the number of features. If there are only two input characteristics, the hyperplane is merely a line. The hyperplane becomes a two-dimensional plane when the number of input features reaches three. When the number of features exceeds three, it becomes impossible to imagine.





4.3 KNN

The K-Nearest Neighbour algorithm is based on the Supervised Learning technique and is one of the most basic Machine Learning algorithms. The K-NN method assumes that the new case/data and existing cases are similar and places the new case in the category that is most similar to the existing categories. The K-NN method stores all available data and classifies a new data point based on its similarity to the existing data. This means that new data can be quickly sorted into a well-defined category using the K-NN method. The K-NN approach can be used for both regression and classification, but it is more commonly utilised for classification tasks. The K-NN algorithm is a non-parametric algorithm, which means it makes no assumptions about the underlying data. It's also known as a lazy learner algorithm since it doesn't learn from the training set right away; instead, it saves the dataset and performs an action on it when it comes time to classify it. During the training phase, the KNN algorithm simply points on either side of the hyperplane. The hyperplane stores the dataset, and when it receives new data, it classifies it into a category that is quite similar to the new data. Consider the following scenario: We have an image of a creature that looks like a cat or a dog, but we don't know whether it's a cat or a dog. We can utilise the KNN method for this identification because it is based on a similarity measure. Our KNN model will look for similarities between the new data set and the cats and dogs photos, and categorise it as either a cat or a dog based on the most similar attributes.

V. SOFTWARE USED

Spyder is a cross-platform open source integrated development environment (IDE) for scientific Python programming. Spyder works with a variety of popular Python packages, including NumPy, SciPy, Matplotlib, pandas, IPython, SymPy, and Cython, as well as other open source applications. It's made available under the MIT licence. Spyder was originally established and developed by Pierre Raybaut in 2009, but it has been maintained and improved by a team of scientific Python developers and the community since 2012.

Features include:

- Syntax highlighting, introspection, and code completion are all features of this editor.
- Multiple IPython consoles are supported.
- The ability to use a GUI to explore and edit variables · A Help window that can automatically or on-demand retrieve and render rich text documentation on functions, classes, and methods.
- IPdb-linked debugger for step-by-step execution
- A code benchmarking tool called a run-time profiler.
- Project support, which allows you to work on numerous development projects at the same time.
- For dealing with the filesystem and managing projects, there's a built-in file explorer.
- A "Find in Files" function that allows for comprehensive regular expression searches over a given range of files.
- A help browser that allows users to search and view Python and package documentation from within the IDE.
- Every user command entered in each console is recorded in a history log.
- An internal console that allows Spyder to examine and regulate its own operations.

VI. CONCLUSIONS

Spyder is a cross-platform open source integrated development environment (IDE) for scientific Python programming. Spyder works with a variety of popular Python packages, including NumPy, SciPy, Matplotlib, pandas, IPython, SymPy, and Cython, as well as other open source applications. It's made available under the MIT licence. Spyder was originally established and developed by Pierre Raybaut in 2009, but it has been maintained and improved by a team of scientific Python developers and the community since 2012.

Features include:

- Syntax highlighting, introspection, and code completion are all features of this editor.
- Multiple IPython consoles are supported.
- The ability to use a GUI to explore and edit variables
- A Help window that can automatically or on-demand retrieve and render rich text documentation on functions, classes, and methods.
- IPdb-linked debugger for step-by-step execution



- Pylint-powered static code analysis
- A code benchmarking tool called a run-time profiler.
- Project support, which allows you to work on numerous development projects at the same time.
- For dealing with the filesystem and managing projects, there's a built-in file explorer.
- A "Find in Files" function that allows for comprehensive regular expression searches over a given range of files.
- A help browser that allows users to search and view Python and package documentation from within the IDE.
- Every user command entered in each console is recorded in a history log.
- An internal console that allows Spyder to examine and regulate its own operations.

REFERENCES

- [1] Kanis JA, Oden A, Johnell O, De Laet C, Jonsson B, Oglesby AK: The components of excess mortality after hip fracture. Bone 32:468–473, (2003)
- [2] Cooper C, Atkinson EJ, Jacobsen SJ, O’Fallon WM, Melton LJ (1993) A population-based study of survival after osteoporotic fractures. Am J Epidemiol 137:1001–1005
- [3] Johnell O, Kanis JA: An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos 17, 1726–1733, (2006)
- [4] Kanis JA, Burlet N, Cooper C, Delmas PD, Reginster JY, Borgstrom F, Rizzoli R: European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int. 2008 Apr;19(4):399-428



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details