

Ensemble Approach for Predicting the Diagnosis of Osteoarthritis Using Soft Voting Classifier

1. Author name and academic degree

Jun-hee Kim (JH Kim), PhD, PT (kemakjh@naver.com)

2. Author's institutional affiliation

Sleep&Spine, Research & Development, KOREATECH Corporation

3. Corresponding author:

Jun-hee Kim, PhD, PT, Senior Researcher

Address: 12 Bongeunsa-ro 49-gil, Gangnam-gu, Seoul, Republic of Korea (06103), Spine&Sleep,
Research & Development, KOREATECH Corporation.

TEL: 82-2-3409-2104

E-mail: kemakjh@naver.com

Abstract

Background: Osteoarthritis (OA) is a common degenerative disease of the joints. Risk factors for OA include non-modifiable factors such as age and gender and modifiable factors such as physical activity.

Purpose: This study aimed to construct a soft voting ensemble model to predict OA diagnosis using variables related to individual characteristics and physical activity and to identify important variables in constructing the model through permutation importance.

Method: Using the RFECV technique, the variables with the best predictive performance were selected among variables, and an ensemble model combining the RandomForest, XGBoost, and LightGBM algorithms was constructed, and the predictive performance and permutation importance of each variable were evaluated.

Result: The variables selected to construct the model were age, gender, grip strength, and quality of life, and the accuracy of the ensemble model was 0.828. The most important variable in constructing the model was age (0.199), followed by grip strength (0.053), quality of life (0.043), and gender (0.034).

Conclusion: The performance of the model for predicting OA was relatively good, and if this model is continuously used and updated, this model could readily be used to predict OA diagnosis and the predictive performance of OA may be further improved.

1 **Introduction**

2 Osteoarthritis (OA) is a common degenerative disease of joints and surrounding structures,
3 and it is estimated that there approximately 300 million patients worldwide suffer from OA.¹
4 Additionally, due to medical expenses and income loss due to OA, social losses of more than
5 303 billion dollars are generated annually in the United States.² For individuals, pain or loss
6 of function due to OA can reduce activities of daily living, which can limit social
7 participation as well as reduce quality of life.³

8 Since OA has such a high prevalence, many studies have been conducted on risk factors
9 related to OA. Risk factors associated with OA, like other diseases, include modifiable and
10 unmodifiable factors.⁴ Non-modifiable risk factors associated with developing OA include
11 age and gender.⁴ Aging is considered the greatest risk factor for developing OA.⁵ The
12 incidence of OA begins to increase rapidly from the age of 50 and stops or decreases after the
13 age of 70.⁶ Gender is also a nonmodifiable factor in the incidence of OA, and the relative risk
14 of hand, knee, and hip OA was found to be 1.52 times greater in women than in men.⁷
15 However, this tendency may differ with age. After the age of 50, the risk of developing OA is
16 higher in women than in men, but OA may be more common in men than in women under the
17 age of 50.⁸

18 Body composition, including height and weight, and physical activity factors, including
19 occupational and recreational activities, can be viewed as modifiable risk factors for
20 developing OA.⁴ Obesity can be seen as a risk factor for developing OA, and the risk of
21 developing OA in obese or overweight people was 2.96 times higher than in people with
22 normal weight.⁹ Additionally, occupational groups with excessive physical activity, such as
23 frequent sitting and standing up, frequent use of joints, or repetitive lifting of objects, have
24 been reported to have a high risk of OA.^{10,11} When repetitive loads are applied to joints by

25 participation in leisure activities such as sports, OA can occur even in young people due to
26 damage to joints and surrounding structures.¹² In particular, those who had experience
27 participating in sports with a high rate of physical contact and high trauma potential, such as
28 soccer, showed a higher rate of OA when they became older.^{12,13}

29 Machine learning is a technique of predicting new data based on attributes learned from a
30 large amount of data.¹⁴ Machine learning shows great performance in estimating values and
31 predicting classifications based on training data.¹⁴ Recently, studies using machine learning to
32 predict the risk of disease or to increase the accuracy of diagnosis using large amounts of
33 medical data are actively being conducted in the medical field.¹⁵ With the development of
34 machine learning techniques in the 2000s, models using algorithms such as support vector
35 machine (SVM) and principal component analysis (PCA) have begun to be applied to knee
36 OA research.¹⁶ These studies mainly used medical images such as X-ray or MRI, or
37 biomechanical data such as kinematics including joint angles or accelerations.¹⁶ Various
38 algorithms are developed and used to build machine learning models, and new algorithms are
39 continuously updated. For example, the XGBoost developed by Chen (2015) and the
40 LightGBM algorithm proposed by Ke (2017) have been developed and actively used as
41 boosting techniques to exhibit higher performance.^{17,18} Additionally, machine learning can be
42 implemented by combining the previously developed techniques, such as ensembles, to
43 construct new models and use them to predict data.¹⁹ The soft voting ensemble technique is a
44 model merging technique that makes predictions through majority voting based on the
45 prediction probabilities of different machine learning models.²⁰ The soft voting ensemble
46 technique can provide better overall results than other base models.²⁰

47 Although a machine learning model that predicts the presence of OA in a specific joint
48 using X-ray or MRI image or biochemical data has been created, a predictive model for OA
49 diagnosis using general information such as individual characteristics including age, gender,

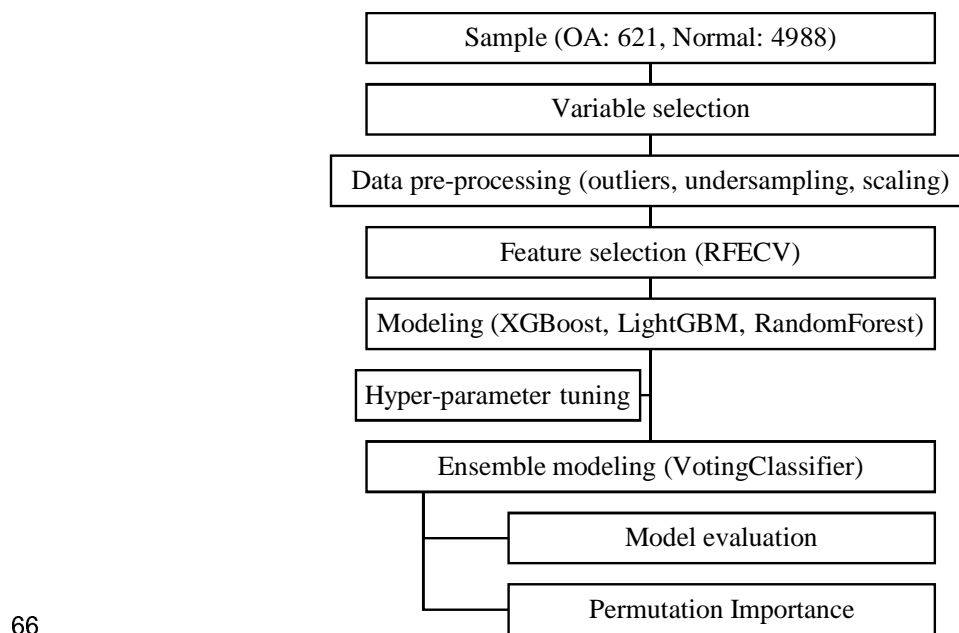
50 height, and weight, and physical activity-related variables including occupational status and
51 leisure activities has not been studied. Therefore, this study aimed to construct a soft voting
52 ensemble model to predict OA diagnosis by using data such as age, gender, individual
53 characteristics, and physical activities of Koreans in the KNHANES (The Korea National
54 Health and Nutrition Examination Survey) database. Additionally, based on the constructed
55 ensemble model, permutation importance is analyzed to identify variables with high
56 importance in predicting OA diagnosis.

57

58 **Methods**

59 1. Study design and population

60 The study samples were obtained from the KNHANES, a national periodic statistic
61 conducted by the Korea Disease Control and Prevention Agency. This study used data from
62 the first year of the 8th KNHANES (2019-2020). Among the 8110 datasets, only samples
63 aged 19 years or older were selected. A total of 621 samples diagnosed with OA but not
64 diagnosed with rheumatoid arthritis and 4988 samples with neither diagnosed OA nor
65 rheumatoid arthritis were used to construct the model.



67 Figure 1. The flowchart diagram of the research process.

68

69 2. Variable selection

70 The variables used in this study, variables representing individual characteristics and
 71 variables related to physical activity were selected. The variables and explanations selected in
 72 this study are shown in Table 1.

Table 1. List of variables and variable descriptions

Variable	OA (621)	Normal (4988)	Description
Sex	M:116 / F:505	M:2403 / F:2585	Gender (M: male / F: female)
Age	67.84 ± 9.95	49.04 ± 16.21	Age (year)
Height	156.72 ± 8.27	164.74 ± 9.06	Height (cm)
Weight	61.35 ± 10.70	64.95 ± 12.97	Weight (kg)
BMI	24.92 ± 3.46	23.82 ± 3.61	Weight/Height ²
BO1_1	1:426 / 2:90 / 3:105	1:3122 / 2:671 / 3:1195	Weight change (1: none / 2: increase / 3: decrease)
HE_wc	87.69 ± 9.46	83.51 ± 10.48	Waist circumference (cm)
GS	21.97 ± 7.68	29.53 ± 9.91	Grip strength (kg)
LQ4_00	1:131 / 2:490	1:278 / 2:4710	Activity limitation (1: yes / 2: no)
EQ5D	0.86 ± 0.15	0.96 ± 0.09	Quality of life (EQ5D index)
EC1_1	1:245 / 2:376	1:3199 / 2:1789	Work (1: yes / 2: no)
EC_wht_23	14.91 ± 21.17	27.91 ± 21.92	Work time
BE3_71	1:2 / 2:619	1:82 / 2:4906	Work activity (1: yes / 2: no)
BE3_75	1:11 / 2:610	1:558 / 2:4430	Leisure activity (1: yes / 2: no)
BP16_1, BP16_2	6.39 ± 1.56	6.88 ± 1.28	Sleep time (hour)

73

74

75 3. Data preprocessing

76 Boxplot was used to remove outliers from numerical data based on the interquartile range
77 (IQR). The ratio of the number of samples of patients with OA to the number of samples of
78 normal subjects was about 1:8, indicating an imbalance in the data distribution. To address
79 the data imbalance, this study used the RandomUnderSampler function, which randomly
80 selects the same number of normal samples as samples diagnosed with OA. Categorical
81 variables were transformed into dummy and numerical variables were scaled with the
82 StandardScaler function. The preprocessed data were classified into training and test data at a
83 ratio of 7:3.

84

85 4. Feature selection

86 The machine learning algorithms used in this study were RandomForest, XGBoost, and
87 LightGBM. In constructing these models, RFECV (Recursive Feature Elimination with Cross
88 Validation) module was used to remove unnecessary variables. Since the same number of
89 features was required to create an ensemble model by combining the three models, the
90 features selected from the model with the best performance among the three models were
91 used to construct the ensemble model.

92

93 5. Machine learning modeling

94 The data was divided into a training data set and a validation data set (hold-out validation),
95 and models of the RandomForest, XGBoost, and LightGBM were built using the training
96 data set, respectively. GridSearchCV was used to find the optimal hyperparameters for each
97 model, and the optimal hyperparameters were set for each model. To improve the accuracy of
98 OA diagnosis prediction, modeling was performed using the soft voting ensemble technique.
99 In this study, an ensemble model was constructed with RandomForest, XGBoost, and
100 LightGBM classifiers to predict OA diagnosis.

101

102 6. Evaluation

103 The performance of the model was evaluated by classifying the predicted and actual
104 classes into a confusion matrix in the form of TP (true positive), FP (false positive), FN (false
105 negative), and TN (true negative) using test data. Models were evaluated by accuracy,
106 precision, recall, F1-score and ROC AUC score calculated based on the confusion matrix.
107 The permutation importance was calculated to identify important variables in the model.
108 Permutation importance indicates the relative importance of variables in construction a model.

109 Results

110 1. Feature selection (RFECV)

111 As a result of performing the RFECV technique, the XGBoost model showed the best
112 performance based on the smallest number of variables (Table 2). The features selected when
113 constructing the ensemble model were Gender, Age, Quality of life, and Grip strength
114 variables based on XGBoost, which showed the highest performance.

Table 2. Feature selection and mean test score of the three models

Model	Feature selection (count)	Mean test score
RandomForest	Gender, Age, Height, Weight, Waist circumference, BMI, Quality of life, Work time, Sleep time, Weight change, Grip strength (11)	0.774
XGBoost	Gender, Age, Quality of life, Grip strength (4)	0.779
LightGBM	Gender, Age, Height, Weight, Waist circumference, BMI, Quality of life, Work, Work time, Leisure activity, Sleep time, Weight change, Grip strength (13)	0.756

115

116 2. Model evaluation

117 The confusion matrix of the ensemble model combining the three models using the
118 validation data set is shown in Table 3. The accuracy score, an overall performance indicator
119 calculated through the confusion matrix, was 0.828. The recall score, which indicates how
120 well the model can identify actual positives, was 0.882. Additionally, the F-1 score was 0.837
121 and the ROC AUC score was 0.829.

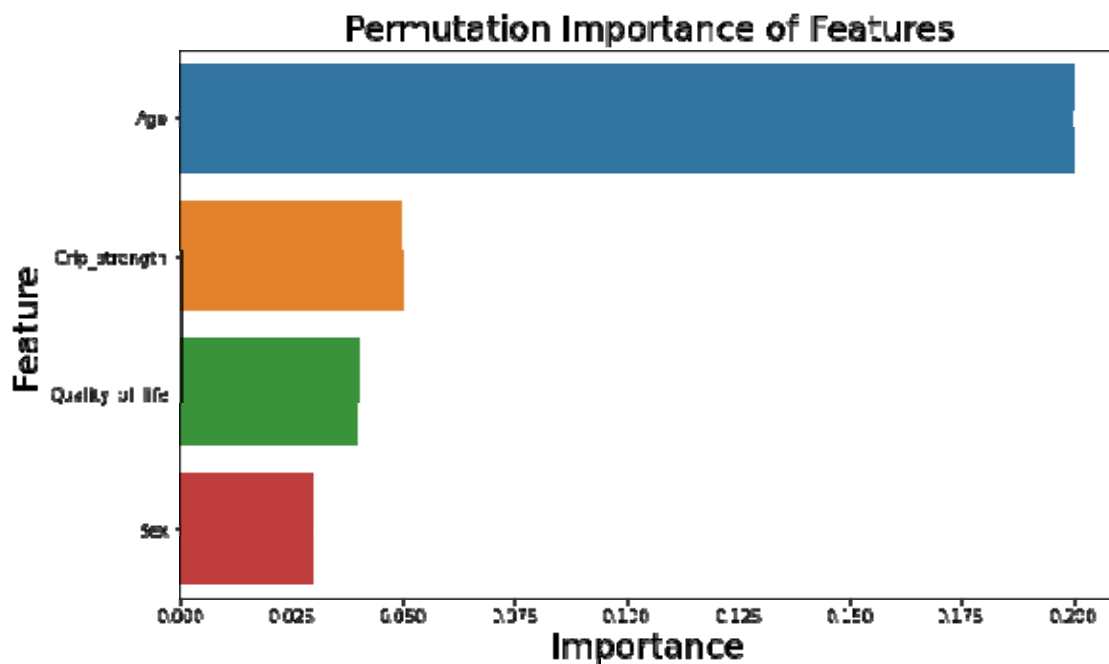
Table 3. Confusion matrix of voting ensemble classification

	Predicted	Positive	Negative
Actual			
Positive		164	22
Negative		42	145

122

123 3. Permutation Importance

124 Figure 2 shows the permutation importance of the ensemble model for predicting OA
125 diagnosis. The most important variable in constructing the model was age, which was 0.199.
126 The second important variable was grip strength, with a significance score of 0.053. The
127 importance of quality of life was 0.043, and gender was 0.034.



128

129 Figure 2. Permutation importance of features in the model

130

131 Discussion

132 In this study, a machine learning classification model was constructed to predict OA
133 diagnosis based on factors related to individual characteristics and physical activities during
134 daily life. The variables selected to build the model were age, gender, grip strength, and
135 quality of life. This classification model was created by combining the latest bagging and
136 boosting classifiers such as RandomForest, XGBoost, and LightGBM. The classification
137 model composed of these ensemble techniques had an accuracy of over 0.8 in predicting OA
138 diagnosis. The variable with the highest importance in constructing the model was age,
139 followed by grip strength, quality of life, and gender in order of importance.

140 In this study, the non-modifiable factors selected as variables in constructing the ensemble
141 model were age and gender. OA is considered a degenerative joint disease, and its prevalence
142 is reported to increase with age.²¹⁻²³ Furthermore, studies have shown that age is the most
143 prominent risk factor for the onset and progression of primary OA in joints such as the
144 interphalangeal, hip and knee.^{7,22,23} OA can be pathologically characterized by irregularly
145 distributed cartilage loss in areas of increased load or synovial inflammation.²⁴ With aging,
146 articular joints including cartilage, subchondral bone, muscle, soft tissue, synovial membrane,
147 and synovial fluid may be deformed, resulting in OA.²⁵ Based on these results, age would
148 have shown the highest permutation importance in constructing an ensemble model in this
149 study. It showed about 4-times higher permutation importance compared to other variables.
150 Additionally, older age and female gender were risk factors for OA.²⁶ Studies of OA and
151 gender have reported a higher prevalence of OA in women than in men.^{26,27} The prevalence
152 of symptomatic radiographic knee OA was 4.6% higher in women than men in one large
153 cohort study, and 5.2% higher in women than men in another cohort study.^{28,29} Additionally,
154 compared with men, women with OA have more severe radiological findings and
155 symptoms.²⁷ However, the explanation for the difference in OA incidence between men and

156 women is still insufficient. It is also unclear whether this is due to differences in the roles of
157 sex hormones and reproductive factors, or whether it is due to weight or obesity, which can
158 be influenced by hormonal factors.^{30,31} Therefore, although the prevalence rate between men
159 and women shows a difference, it may not be a variable that can completely distinguish the
160 diagnosis of OA because the explanation for the difference is not sufficient.^{30,31} For this
161 reason, the gender variable was selected to generate the ensemble model, but it showed the
162 lowest permutation importance among the variables.

163 The modifiable factors selected as variables in constructing the ensemble model were grip
164 strength and quality of life. Previous studies to identify the relationship between grip strength
165 and OA have been continuously conducted.⁵ These studies used grip strength as a variable
166 and found that people with OA of the hand mainly had reduced grip strength compared to the
167 general population, and it was thought that the risk of OA could be predicted through grip
168 strength.³²⁻³⁴ Based on these results, it was thought that improving grip strength through
169 training would help improve OA symptoms.³⁵ However, grip strength is not only a marker for
170 OA but also a marker for general health, and since grip strength can vary according to gender
171 or physical characteristics, it was considered difficult to determine OA of the hand only with
172 grip strength.³² Therefore, grip strength was also a factor adopted as a variable in constructing
173 the model, but it would have shown permutation importance similar to other variables except
174 age. The EQ-5D used to measure quality of life in this study consists of 5 items that measure
175 motor skills, self-management, activities of daily living, pain/discomfort, anxiety/depression
176 in 3 levels and 1 item that measures health status in a visual analog scale.³⁶ Compared to
177 assessment tools that measure the health status of OA patients, such as the Western Ontario
178 and MacMaster Universities (WOMAC) OA index, the EQ-5D showed overall good validity
179 and reliability in assessing the health status of patients with OA.^{37,38} Also, the EQ-5D has
180 fewer and simpler questions, so it may be more appropriate for elderly subjects.^{38,39} Therefore,

181 the EQ-5D is used to evaluate the overall health status of relatively elderly patients with
182 OA.^{40,41} For this reason, the quality of life variable evaluated by EQ-5D would have been
183 selected in constructing the ensemble model. Similarly, since the quality of life can be
184 influenced by various other factors such as psychology, physical condition, and social
185 environment, it would have shown low permutation importance compared with other
186 variables.

187 In the KNee OsteoArthritis Prediction challenge held in 2020 (KNOAP2020), various
188 machine learning models for predicting knee OA have been submitted using X-ray and MRI
189 image and clinical data such as age, BMI, affected side, and history of knee injury.⁴² The
190 model that showed the highest ROC AUC score in the competition was a model that applied
191 the logistic regression algorithm using X-ray and clinical data and scored 0.636.⁴² The model
192 with the next highest ROC AUC score was an ensemble model using X-ray, MRI, and clinical
193 data, with a score of 0.624.⁴² According to Hosnijeh (2018), the ROC AUC score of the hip
194 OA model, which was constructed by combining imaging, biochemical, and genetic variables
195 with variables such as age, obesity, waist circumference, and gender, was between 0.67 and
196 0.82.⁴³ Although image data such as X-ray or MRI and biomechanical or genetic factors were
197 not used as variables, the ROC AUC score of the ensemble model constructed in this study
198 was relatively high. The reasons for the relatively high performance of the ensemble model in
199 this study may be the following difference. The data used in KNAOP2020 were obtained
200 from the Osteoarthritis Initiative (OAI) database.⁴² The average age of the subjects in this
201 dataset was over 61 years old.⁴² Additionally, in the Hosnijeh (2018) study, all of the selected
202 subjects were over 55 years of age, and the average age of the samples used to construct the
203 model ranged from 55.8 to 64 years.⁴³ The age of the dataset used in our study was over 19
204 years old, and the average age of subjects with OA was 67 years old, and the average age of
205 normal subjects was 49 years old, showing an age difference of about 18 years. Additionally,

206 while other studies have constructed prediction models for knee, hip, and wrist OA separately,
207 this study constructed a model that included all OA without distinction. In the case of older
208 subjects, the possibility of having OA in one of the body parts would be higher than that of
209 younger subjects.^{5,7,22} Therefore, the age variable in the ensemble model constructed in this
210 study showed great permutation importance in the classification compared with other
211 variables, and it would have shown high classification performance despite the absence of
212 imaging data, genetic, and biochemical data.

213 This study had several limitations. First, the constructed model showed relatively high
214 predictive performance, but the ensemble model based on the machine learning algorithm
215 cannot explain the causal relationship between OA and other variables. Second, because it is
216 not a predictive model for OA of a specific joint, this model can predict the overall risk of
217 OA, but it will be difficult to use to predict or diagnose OA of a specific joint. Based on these
218 limitations, future research will require the development and application of an algorithm that
219 can infer causal relationships, and modeling to predict arthritis of a specific joint will be
220 needed.

221

222 **Conclusion**

223 In this study, a model to predict OA was created using individual characteristics and
224 physical activity-related variables. This model was constructed by selecting variables of age,
225 gender, grip strength, and quality of life, and was composed of an ensemble model combining
226 RandomForest, XGBoost, and LightGBM models with high predictive performance. The
227 performance of the ensemble model for predicting OA diagnosis was relatively good, and this
228 ensemble model can be used to easily predict whether a person will be diagnosed with OA.
229 Additionally, if the model is updated by continuously accumulating new data, the prediction
230 performance can be further improved.

231

232 **Acknowledgments**

233 This research did not receive any specific grant from funding agencies in the public,
234 commercial, or not-for-profit sectors. This study was conducted using de-identified data from
235 the 8th KNHANES performed by the Korea Disease Control and Prevention Agency in 2019-
236 2020.

237

238

239 **References**

- 240 1. James SL, Abate D, Abate KH, et al. Global, regional, and national incidence,
241 prevalence, and years lived with disability for 354 diseases and injuries for 195
242 countries and territories, 1990–2017: a systematic analysis for the Global Burden of
243 Disease Study 2017. *Lancet*. 2018;392(10159):1789-1858.
- 244 2. Murphy LB, Cisternas MG, Pasta DJ, Helmick CG, Yelin EH. Medical expenditures
245 and earnings losses among US adults with arthritis in 2013. *Arthritis Care Res*.
246 2018;70(6):869-876.
- 247 3. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of
248 osteoarthritis. *Nat Rev Rheumatol*. 2014;10(7):437-441.
- 249 4. Abramoff B, Caldera FE. Osteoarthritis: pathology, diagnosis, and treatment options.
250 *Med Clin*. 2020;104(2):293-311.
- 251 5. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Clin Geriatr Med*.
252 2010;26(3):355-369.
- 253 6. Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic
254 hand, hip, and knee osteoarthritis among patients in a health maintenance organization.
255 *Arthritis Rheum Off J Am Coll Rheumatol*. 1995;38(8):1134-1141.
- 256 7. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK.
257 Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis:
258 influences of age, gender and osteoarthritis affecting other joints. *Ann Rheum Dis*.
259 2014;73(9):1659-1664.
- 260 8. Felson DT, Zhang Y, Hannan MT, et al. The incidence and natural history of knee
261 osteoarthritis in the elderly, the framingham osteoarthritis study. *Arthritis Rheum*.

- 262 1995;38(10):1500-1505.
- 263 9. Blagojevic M, Jinks C, Jeffery A, Jordan IKP. Risk factors for onset of osteoarthritis
264 of the knee in older adults: a systematic review and meta-analysis. *Osteoarthr Cartil.*
265 2010;18(1):24-33.
- 266 10. Gignac MAM, Irvin E, Cullen K, et al. Men and women's occupational activities and
267 the risk of developing osteoarthritis of the knee, hip, or hands: a systematic review and
268 recommendations for future research. *Arthritis Care Res.* 2020;72(3):378-396.
- 269 11. Muraki S, Akune T, Oka H, et al. Association of occupational activity with
270 radiographic knee osteoarthritis and lumbar spondylosis in elderly patients of
271 population-based cohorts: a large-scale population-based study. *Arthritis Care Res*
272 *Off J Am Coll Rheumatol.* 2009;61(6):779-786.
- 273 12. Buckwalter JA, Lane NE. Athletics and osteoarthritis. *Am J Sports Med.*
274 1997;25(6):873-881.
- 275 13. Prien A, Boudabous S, Junge A, Verhagen E, Delattre B, Tscholl PM. Every second
276 retired elite female football player has MRI evidence of knee osteoarthritis before age
277 50 years: a cross-sectional study of clinical and MRI outcomes. *Knee Surgery, Sport*
278 *Traumatol Arthrosc.* 2020;28(2):353-362.
- 279 14. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects.
280 *Science (80-).* 2015;349(6245):255-260.
- 281 15. Fatima M, Pasha M. Survey of machine learning algorithms for disease diagnostic. *J*
282 *Intell Learn Syst Appl.* 2017;9(01):1.
- 283 16. Kokkotis C, Moustakidis S, Papageorgiou E, Giakas G, Tsaopoulos DE. Machine
284 learning in knee osteoarthritis: a review. *Osteoarthr Cartil Open.* 2020;2(3):100069.

- 285 17. Chen T, He T, Benesty M, et al. Xgboost: extreme gradient boosting. *R Packag*
286 *version 04-2*. 2015;1(4):1-4.
- 287 18. Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision
288 tree. *Adv Neural Inf Process Syst*. 2017;30.
- 289 19. Massonis G, Villaverde AF, Banga JR. Improving dynamic predictions with ensembles
290 of observable models. *Bioinformatics*. 2023;39(1):btac755.
- 291 20. Islam R, Shahjalal MA. Soft voting-based ensemble approach to predict early stage drc
292 violations. In: *2019 IEEE 62nd International Midwest Symposium on Circuits and*
293 *Systems (MWSCAS)*. IEEE; 2019:1081-1084.
- 294 21. Anderson AS, Loeser RF. Why is osteoarthritis an age-related disease? *Best Pract Res*
295 *Clin Rheumatol*. 2010;24(1):15-26.
- 296 22. Loeser RF. Age-related changes in the musculoskeletal system and the development of
297 osteoarthritis. *Clin Geriatr Med*. 2010;26(3):371-386.
- 298 23. Li Y, Wei X, Zhou J, Wei L. The age-related changes in cartilage and osteoarthritis.
299 *Biomed Res Int*. 2013;2013.
- 300 24. Mankin HJ, Brandt KD, Shulman LE. Workshop on etiopathogenesis of osteoarthritis,
301 Warrenton VA, July 21-25, 1985: reasearch recommendations arising from the
302 workshop. *J Rheumatol*. 1986;13(6):1126-1160.
- 303 25. Hamerman D. Biology of the aging joint. *Clin Geriatr Med*. 1998;14(3):417-434.
- 304 26. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis:
305 a review. *Jama*. 2021;325(6):568-578.
- 306 27. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis

- 307 of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthr*
308 *Cartil.* 2005;13(9):769-781.
- 309 28. Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The prevalence
310 of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis*
311 *Rheum Off J Am Coll Rheumatol.* 1987;30(8):914-918.
- 312 29. Jordan JM, Helmick CG, Renner JB, et al. Prevalence of knee symptoms and
313 radiographic and symptomatic knee osteoarthritis in African Americans and
314 Caucasians: the Johnston County Osteoarthritis Project. *J Rheumatol.* 2007;34(1):172-
315 180.
- 316 30. de Klerk BM, Schiphof D, Groeneveld FPMJ, et al. No clear association between
317 female hormonal aspects and osteoarthritis of the hand, hip and knee: a systematic
318 review. *Rheumatology.* 2009;48(9):1160-1165.
- 319 31. Hussain SM, Cicuttini FM, Alyousef B, Wang Y. Female hormonal factors and
320 osteoarthritis of the knee, hip and hand: a narrative review. *Climacteric.*
321 2018;21(2):132-139.
- 322 32. Haugen IK, Aaserud J, Kvien TK. Get a grip on factors related to grip strength in
323 persons with hand osteoarthritis: results from an observational cohort study. *Arthritis*
324 *Care Res.* 2021;73(6):794-800.
- 325 33. Dominick KL, Jordan JM, Renner JB, Kraus VB. Relationship of radiographic and
326 clinical variables to pinch and grip strength among individuals with osteoarthritis.
327 *Arthritis Rheum.* 2005;52(5):1424-1430.
- 328 34. Allen KD, Jordan JM, Renner JB, Kraus VB. Relationship of global assessment of
329 change to AUSCAN and pinch and grip strength among individuals with hand

- 330 osteoarthritis. *Osteoarthr Cartil.* 2006;14(12):1281-1287.
- 331 35. Rogers MW, Wilder F V. The effects of strength training among persons with hand
332 osteoarthritis: a two-year follow-up study. *J hand Ther.* 2007;20(3):244-250.
- 333 36. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. *Appl*
334 *Health Econ Health Policy.* 2017;15(2):127-137.
- 335 37. Roos LS Lohmander, EM MK. WOMAC Osteoarthritis Index: Reliability, validity,
336 and responsiveness in patients with arthroscopically assessed osteoarthritis. *Scand J*
337 *Rheumatol.* 1999;28(4):210-215.
- 338 38. Lim N-Y, Lee I, Lee E-N, et al. A validation study of EQ-5D in the patients with
339 osteoarthritis. *J muscle Jt Heal.* 2010;17(2):203-211.
- 340 39. Wailoo A, Hernandez Alava M, Escobar Martinez A. Modelling the relationship
341 between the WOMAC osteoarthritis index and EQ-5D. *Health Qual Life Outcomes.*
342 2014;12(1):1-6.
- 343 40. García-Pérez L, Ramos-García V, Serrano-Aguilar P, et al. EQ-5D-5L utilities per
344 health states in Spanish population with knee or hip osteoarthritis. *Health Qual Life*
345 *Outcomes.* 2019;17(1):1-14.
- 346 41. Bilbao A, García-Pérez L, Arenaza JC, et al. Psychometric properties of the EQ-5D-5L
347 in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual*
348 *Life Res.* 2018;27(11):2897-2908.
- 349 42. Hirvasniemi J, Runhaar J, van der Heijden RA, et al. The KNee OsteoArthritis
350 Prediction (KNOAP2020) challenge: An image analysis challenge to predict incident
351 symptomatic radiographic knee osteoarthritis from MRI and X-ray images. *Osteoarthr*
352 *Cartil.* 2023;31(1):115-125.

353 43. Hosnijeh FS, Kavousi M, Boer CG, et al. Development of a prediction model for
354 future risk of radiographic hip osteoarthritis. *Osteoarthr Cartil.* 2018;26(4):540-546.

355