

Supplementary Table S1: The clinical characteristics of participants.

Characteristics	Rheumatoid arthritis	Osteoarthritis
Number	25	25
Age (years) <sup>a</sup>	63.25 ± 10.24	65.34 ± 13.25
Sex (Male/Female)	5/20	5/20
Anti-CCP (IU/mL) <sup>b</sup>	102.50 (37.00, 251.50)	/
RF (IU/mL) <sup>b</sup>	67.00 (35.50, 148.50)	/
CRP (mm/h) <sup>b</sup>	10.35 (3.16, 17.78)	/
ESR (mg/mL) <sup>b</sup>	31.00 (24.50, 41.00)	/

<sup>a</sup> Expressed as the median (25th to 75th percentile).

<sup>b</sup> Expressed as the means ± SD (standard deviation).

Abbreviations: anti-CCP, anti-cyclic citrullinated peptide antibody; RF, rheumatoid factor; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate.







Supplementary Table S4: The mean result of 10 single runs of the machine learning (ML) model for the training set in RA vs. HC.

ML model	AUC(95%CI)	cutof(95%CI)	Accuracy(95%CI)	Sensitivity(95%CI)	Specificity(95%CI)	Positive predictive value(95%CI)	Negative predictive value(95%CI)	F1 score(95%CI)	Kappa(95%CI)
XGBoost	0.999 (0.997-1.000)	0.906(0.877-0.934)	0.975(0.969-0.981)	0.979(0.972-0.985)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.864(0.837-0.892)	0.989(0.986-0.993)	0.912(0.892-0.931)
logistic	0.988 (0.972-1.000)	0.664(0.616-0.712)	0.963(0.955-0.972)	0.972(0.962-0.982)	0.959(0.949-0.969)	0.992(0.991-0.994)	0.833(0.807-0.858)	0.982(0.977-0.987)	0.869(0.848-0.890)
LightGBM	1.000 (NaN-1.000)	0.873(0.832-0.915)	0.993(0.993-0.993)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.958(0.956-0.960)	1.000(1.000-1.000)	0.974(0.973-0.975)
RandomForest	1.000 (NaN-1.000)	0.700(0.631-0.769)	0.988(0.983-0.992)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.926(0.897-0.954)	1.000(1.000-1.000)	0.953(0.935-0.972)
AdaBoost	1.000 (NaN-1.000)	0.695(0.660-0.730)	0.992(0.990-0.993)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.949(0.937-0.961)	1.000(1.000-1.000)	0.969(0.961-0.976)
CNB	0.744 (0.667-0.821)	0.747(0.682-0.812)	0.676(0.655-0.698)	0.629(0.599-0.660)	0.969(0.942-0.995)	0.992(0.985-0.999)	0.326(0.313-0.339)	0.769(0.748-0.790)	0.328(0.311-0.345)
MLP	0.816 (0.781-0.849)	0.915(0.773-1.057)	0.790(0.617-0.963)	0.764(0.554-0.975)	0.968(0.949-0.988)	NaN(NaN-NaN)	0.585(0.435-0.736)	NaN(NaN-NaN)	0.607(0.412-0.802)
SVM	0.980 (0.952-1.000)	0.610(0.569-0.651)	0.962(0.958-0.965)	0.970(0.965-0.975)	0.961(0.952-0.970)	0.992(0.991-0.994)	0.829(0.812-0.847)	0.981(0.979-0.983)	0.867(0.857-0.877)
KNN	0.992 (0.982-1.000)	0.840(0.762-0.918)	0.691(0.460-0.922)	0.947(0.933-0.962)	0.975(0.961-0.988)	NaN(NaN-NaN)	0.524(0.359-0.689)	NaN(NaN-NaN)	0.533(0.303-0.763)

Supplementary Table S5: The mean result of 10 single runs of the machine learning (ML) model for the testing set in RA vs. HC.

ML model	AUC(95%CI)	cutof(95%CI)	Accuracy(95%CI)	Sensitivity(95%CI)	Specificity(95%CI)	Positive predictive value(95%CI)	Negative predictive value(95%CI)	F1 score(95%CI)	Kappa(95%CI)
XGBoost	0.972 (NaN-NaN)	0.906(0.877-0.934)	0.950(0.934-0.966)	0.974(0.958-0.990)	0.975(0.926-1.024)	0.993(0.985-1.002)	0.754(0.663-0.845)	0.983(0.974-0.993)	0.807(0.732-0.882)
logistic	0.989 (NaN-NaN)	0.664(0.616-0.712)	0.958(0.942-0.975)	0.972(0.953-0.992)	0.978(0.934-1.021)	0.983(0.967-0.998)	0.851(0.768-0.934)	0.977(0.964-0.990)	0.856(0.797-0.916)
LightGBM	0.966 (NaN-NaN)	0.873(0.832-0.915)	0.925(0.896-0.954)	0.925(0.884-0.966)	0.975(0.942-1.008)	0.973(0.956-0.989)	0.737(0.619-0.854)	0.948(0.920-0.976)	0.722(0.632-0.811)
RandomForest	0.979 (NaN-NaN)	0.700(0.631-0.769)	0.922(0.883-0.961)	0.967(0.939-0.996)	0.980(0.941-1.019)	0.983(0.969-0.998)	0.729(0.589-0.869)	0.975(0.954-0.996)	0.743(0.617-0.869)
AdaBoost	0.983 (NaN-NaN)	0.695(0.660-0.730)	0.897(0.862-0.933)	0.952(0.904-1.001)	1.000(1.000-1.000)	0.989(0.974-1.004)	0.600(0.478-0.723)	0.970(0.937-1.002)	0.662(0.549-0.774)
CNB	0.755 (0.601-0.908)	0.747(0.682-0.812)	0.678(0.624-0.731)	0.667(0.600-0.734)	1.000(1.000-1.000)	0.972(0.952-0.993)	0.302(0.255-0.349)	0.787(0.739-0.835)	0.293(0.231-0.356)
MLP	0.808 (NaN-NaN)	0.915(0.773-1.057)	0.786(0.590-0.982)	0.766(0.540-0.992)	0.971(0.934-1.009)	NaN(NaN-NaN)	0.662(0.467-0.857)	NaN(NaN-NaN)	0.615(0.399-0.831)
SVM	0.983 (NaN-NaN)	0.610(0.569-0.651)	0.956(0.941-0.970)	0.974(0.957-0.990)	0.983(0.951-1.016)	0.987(0.976-0.997)	0.799(0.729-0.868)	0.980(0.969-0.991)	0.825(0.787-0.864)
KNN	0.974 (NaN-NaN)	0.840(0.762-0.918)	0.694(0.471-0.917)	0.968(0.949-0.987)	0.969(0.928-1.010)	NaN(NaN-NaN)	0.492(0.319-0.664)	NaN(NaN-NaN)	0.489(0.266-0.717)

Supplementary Table S6: The mean result of 10 single runs of the machine learning (ML) model for the training set in RA vs. OA.

ML model	AUC(95%CI)	cutof(95%CI)	Accuracy(95%CI)	Sensitivity(95%CI)	Specificity(95%CI)	Positive predictive value(95%CI)	Negative predictive value(95%CI)	F1 score(95%CI)	Kappa(95%CI)
XGBoost	1.000 (NaN-NaN)	0.606(0.549-0.664)	0.993(0.993-0.993)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.946(0.943-0.949)	1.000(1.000-1.000)	0.968(0.966-0.970)
logistic	0.921 (0.862-0.981)	0.810(0.786-0.835)	0.855(0.842-0.868)	0.865(0.848-0.882)	0.843(0.817-0.869)	0.974(0.970-0.978)	0.466(0.430-0.503)	0.916(0.907-0.924)	0.519(0.485-0.552)
LightGBM	1.000 (NaN-NaN)	0.751(0.693-0.809)	0.993(0.993-0.993)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.946(0.942-0.950)	1.000(1.000-1.000)	0.968(0.966-0.970)
RandomForest	1.000 (NaN-NaN)	0.675(0.616-0.734)	0.987(0.979-0.995)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.913(0.866-0.959)	1.000(1.000-1.000)	0.945(0.913-0.978)
AdaBoost	1.000 (NaN-NaN)	0.541(0.535-0.547)	0.993(0.993-0.993)	1.000(1.000-1.000)	1.000(1.000-1.000)	1.000(1.000-1.000)	0.946(0.943-0.950)	1.000(1.000-1.000)	0.968(0.967-0.970)
CNB	0.532 (0.430-0.634)	0.605(0.532-0.679)	0.473(0.438-0.507)	0.425(0.381-0.468)	0.867(0.825-0.909)	0.957(0.943-0.971)	0.176(0.165-0.188)	0.585(0.541-0.628)	0.106(0.095-0.117)
MLP	0.618 (0.533-0.703)	1.048(0.901-1.195)	0.533(0.349-0.717)	0.487(0.271-0.704)	0.904(0.841-0.968)	NaN(NaN-NaN)	0.251(0.184-0.318)	NaN(NaN-NaN)	0.221(0.114-0.327)
SVM	0.813 (0.690-0.936)	0.853(0.873-0.893)	0.851(0.844-0.858)	0.881(0.875-0.887)	0.702(0.677-0.727)	0.952(0.947-0.958)	0.446(0.432-0.461)	0.915(0.911-0.920)	0.461(0.442-0.480)
KNN	0.939 (0.897-0.980)	0.980(0.941-1.019)	0.191(0.066-0.315)	0.735(0.690-0.781)	0.983(0.951-1.016)	NaN(NaN-NaN)	0.150(0.105-0.194)	NaN(NaN-NaN)	0.041(-0.039-0.121)

Supplementary Table S7: The mean result of 10 single runs of the machine learning (ML) model for the testing set in RA vs. OA.

分类模型	AUC(95%CI)	cutof(95%CI)	准确度(95%CI)	灵敏度(95%CI)	特异度(95%CI)	阳性预测值(95%CI)	阴性预测值(95%CI)	F1分数(95%CI)	Kappa(95%CI)
XGBoost	0.801 (NaN-NaN)	0.606(0.549-0.664)	0.891(0.858-0.925)	0.901(0.844-0.958)	0.801(0.704-0.897)	0.936(0.909-0.963)	NaN(NaN-NaN)	0.917(0.878-0.956)	0.444(0.251-0.636)
logistic	0.935 (0.844-0.998)	0.810(0.786-0.835)	0.840(0.799-0.881)	0.839(0.754-0.924)	1.000(1.000-1.000)	0.978(0.962-0.994)	0.427(0.327-0.528)	0.898(0.835-0.961)	0.474(0.376-0.572)
LightGBM	0.869 (NaN-NaN)	0.751(0.693-0.809)	0.906(0.891-0.920)	0.924(0.889-0.959)	0.874(0.781-0.967)	0.958(0.936-0.979)	0.619(0.471-0.767)	0.939(0.920-0.958)	0.568(0.474-0.661)
RandomForest	0.798 (NaN-NaN)	0.675(0.616-0.734)	0.880(0.831-0.929)	0.808(0.624-0.993)	0.849(0.740-0.959)	0.940(0.911-0.968)	0.540(0.357-0.723)	0.825(0.641-1.010)	0.482(0.291-0.673)
AdaBoost	0.895 (0.748-1.000)	0.541(0.535-0.547)	0.883(0.850-0.916)	0.890(0.812-0.967)	0.878(0.787-0.968)	0.952(0.931-0.973)	0.505(0.404-0.607)	0.915(0.867-0.964)	0.497(0.375-0.618)
CNB	0.482 (NaN-NaN)	0.605(0.532-0.679)	0.406(0.364-0.447)	0.379(0.293-0.465)	0.969(0.926-1.012)	0.907(0.848-0.965)	0.139(0.101-0.176)	0.525(0.436-0.615)	0.022(-0.018-0.061)
MLP	0.611 (NaN-NaN)	1.048(0.901-1.195)	0.534(0.367-0.701)	0.561(0.306-0.816)	0.910(0.791-1.029)	NaN(NaN-NaN)	0.224(0.174-0.274)	NaN(NaN-NaN)	0.166(0.073-0.258)
SVM	0.721 (0.384-0.966)	0.883(0.873-0.893)	0.800(0.772-0.828)	0.874(0.737-1.011)	0.680(0.521-0.839)	0.933(0.906-0.960)	0.330(0.245-0.415)	0.887(0.790-0.984)	0.294(0.198-0.390)
KNN	0.769 (NaN-NaN)	0.980(0.941-1.019)	0.183(0.064-0.302)	0.823(0.764-0.882)	0.707(0.583-0.831)	NaN(NaN-NaN)	0.131(0.100-0.163)	NaN(NaN-NaN)	0.015(-0.014-0.043)