

**Table S7: Accuracies of the 15 machine learning configurations on the testing set.**

	SciSpaCy	MetaMap	Custom	Combine1	Combine2	Combine3	All
<b>Logistic regression</b>	0.677	0.743	0.608	0.75	0.653	0.76	0.764
<b>AdaBoost</b>	0.635	0.753	0.608	0.781	0.681	0.767	0.781
<b>Bagging</b>	0.639	0.736	0.622	0.767	0.66	0.767	0.767
<b>Decision tree</b>	0.604	0.674	0.597	0.674	0.59	0.677	0.684
<b>SVC_rbf</b>	0.656	0.733	0.618	0.722	0.656	0.722	0.722
<b>SVC_poly</b>	0.604	0.625	0.556	0.615	0.604	0.625	0.615
<b>SVC_sigmoid</b>	0.594	0.639	0.556	0.649	0.594	0.656	0.656
<b>kNN</b>	0.615	0.715	0.601	0.719	0.642	0.712	0.715
<b>Naïve bayes</b>	0.59	0.594	0.601	0.604	0.59	0.594	0.604
<b>Random forest</b>	0.684	0.795	0.618	0.778	0.698	0.781	0.788
<b>SGDClassifier_L1</b>	0.583	0.747	0.594	0.726	0.611	0.757	0.715
<b>SGDClassifier_L2</b>	0.569	0.76	0.503	0.75	0.566	0.74	0.747
<b>SGDClassifier_EN</b>	0.576	0.753	0.601	0.733	0.58	0.747	0.743
<b>LinearSVC_L1</b>	0.667	0.743	0.601	0.733	0.663	0.743	0.736
<b>LinearSVC_L2</b>	0.667	0.743	0.601	0.736	0.663	0.747	0.74

AdaBoost: Adaptive Boosting; Bagging: Bootstrap aggregating; SVC\_rbf or \_poly or \_sigmoid: support vector machine with rbf kernel or polynomial kernel or sigmoid kernel;

SGDClassifier\_L1 or \_L2 or \_EN: stochastic gradient descent with L1 or L2 or Elastic Net regularization; NN: k-Nearest Neighbors; LinearSVC\_L1 or \_L2: support vector machine with linear kernel coupled with L1 or L2 regularization

Sci: features extracted by scispaCy; Meta: features extracted by MetaMap; Custom: features extracted by the spaCy customized model; Combine1: features extracted by scispaCy or MetaMap; Combine2: features extracted by scispaCy or the spaCy customized model; Combine3: features extracted by MetaMap or the spaCy customized model; All: features extracted by scispaCy, MetaMap, or the spaCy customized model.

**Table S8: F1 scores of the 15 machine learning configurations on the testing set.**

	SciSpaCy	MetaMap	Custom	Combine1	Combine2	Combine3	All
<b>Logistic regression</b>	0.710	0.752	0.664	0.821	0.693	0.772	0.772
<b>AdaBoost</b>	0.714	0.751	0.62	0.774	0.685	0.765	0.774
<b>Bagging</b>	0.645	0.741	0.64	0.783	0.68	0.779	0.776
<b>Decision tree</b>	0.601	0.674	0.613	0.693	0.614	0.674	0.698
<b>SVC_rbf</b>	0.706	0.72	0.663	0.704	0.71	0.712	0.71
<b>SVC_poly</b>	0.715	0.729	0.692	0.722	0.718	0.729	0.722
<b>SVC_sigmoid</b>	0.651	0.658	0.573	0.664	0.653	0.673	0.673
<b>kNN</b>	0.613	0.723	0.628	0.720	0.633	0.724	0.73
<b>Naïve bayes</b>	0.699	0.693	0.674	0.702	0.697	0.693	0.712
<b>Random forest</b>	0.681	0.794	0.638	0.778	0.695	0.780	0.786
<b>SGDClassifier_L1</b>	0.618	0.767	0.679	0.73	0.654	0.764	0.723
<b>SGDClassifier_L2</b>	0.584	0.751	0.665	0.755	0.542	0.719	0.731
<b>SGDClassifier_EN</b>	0.639	0.756	0.685	0.734	0.623	0.759	0.748

<b>LinearSVC_L1</b>	0.702	0.748	0.657	0.741	0.705	0.752	0.750
<b>LinearSVC_L2</b>	0.702	0.748	0.657	0.744	0.705	0.756	0.750

AdaBoost: Adaptive Boosting; Bagging: Bootstrap aggregating; SVC\_rbf or \_poly or \_sigmoid: support vector machine with rbf kernel or polynomial kernel or sigmoid kernel; SGDClassifier\_L1 or \_L2 or \_EN: stochastic gradient descent with L1 or L2 or Elastic Net regularization; NN: k-Nearest Neighbors; LinearSVC\_L1 or \_L2: support vector machine with linear kernel coupled with L1 or L2 regularization

Sci: features extracted by scispaCy; Meta: features extracted by MetaMap; Custom: features extracted by the spaCy customized model; Combine1: features extracted by scispaCy or MetaMap; Combine2: features extracted by scispaCy or the spaCy customized model; Combine3: features extracted by MetaMap or the spaCy customized model; All: features extracted by scispaCy, MetaMap, or the spaCy customized model.

**Table S9: Recalls of the 15 machine learning configurations on the testing set.**

	SciSpaCy	MetaMap	Custom	Combine1	Combine2	Combine3	All
<b>Logistic regression</b>	0.760	0.747	0.747	0.747	0.753	0.780	0.767
<b>AdaBoost</b>	0.647	0.713	0.649	0.72	0.667	0.727	0.72
<b>Bagging</b>	0.640	0.727	0.647	0.807	0.693	0.787	0.773
<b>Decision tree</b>	0.573	0.647	0.613	0.707	0.627	0.640	0.700
<b>SVC_rbf</b>	0.793	0.660	0.720	0.633	0.707	0.660	0.653
<b>SVC_poly</b>	0.953	0.967	0.960	0.960	0.967	0.967	0.96
<b>SVC_sigmoid</b>	0.727	0.667	0.573	0.667	0.733	0.68	0.673
<b>kNN</b>	0.587	0.713	0.647	0.693	0.593	0.727	0.740
<b>Naïve bayes</b>	0.913	0.880	0.793	0.893	0.907	0.880	0.893
<b>Random forest</b>	0.647	0.760	0.647	0.747	0.660	0.747	0.747

<b>SGDClassifier_L1</b>	0.669	0.828	0.855	0.738	0.731	0.779	0.738
<b>SGDClassifier_L2</b>	0.600	0.717	0.979	0.765	0.510	0.662	0.683
<b>SGDClassifier_EN</b>	0.745	0.759	0.862	0.731	0.690	0.793	0.759
<b>LinearSVC_L1</b>	0.753	0.733	0.733	0.733	0.733	0.747	0.760
<b>LinearSVC_L2</b>	0.753	0.733	0.733	0.74	0.773	0.753	0.760

AdaBoost: Adaptive Boosting; Bagging: Bootstrap aggregating; SVC\_rbf or \_poly or \_sigmoid: support vector machine with rbf kernel or polynomial kernel or sigmoid kernel;

SGDClassifier\_L1 or \_L2 or \_EN: stochastic gradient descent with L1 or L2 or Elastic Net regularization; NN: k-Nearest Neighbors; LinearSVC\_L1 or \_L2: support vector machine with linear kernel coupled with L1 or L2 regularization

Sci: features extracted by scispaCy; Meta: features extracted by MetaMap; Custom: features extracted by the spaCy customized model; Combine1: features extracted by scispaCy or MetaMap; Combine2: features extracted by scispaCy or the spaCy customized model; Combine3: features extracted by MetaMap or the spaCy customized model; All: features extracted by scispaCy, MetaMap, or the spaCy customized model.

**Table S10: Specificities of the 15 machine learning configurations on the testing set.**

	SciSpaCy	MetaMap	Custom	Combine1	Combine2	Combine3	All
<b>Logistic regression</b>	0.678	0.729	0.624	0.732	0.684	0.756	0.75
<b>AdaBoost</b>	0.644	0.754	0.583	0.745	0.646	0.732	0.736
<b>Bagging</b>	0.674	0.769	0.603	0.75	0.655	0.752	0.74
<b>Decision tree</b>	0.579	0.647	0.612	0.648	0.576	0.643	0.672
<b>SVC_rbf</b>	0.646	0.713	0.61	0.673	0.701	0.716	0.679
<b>SVC_poly</b>	0.842	0.875	0.767	0.846	0.829	0.875	0.846
<b>SVC_sigmoid</b>	0.690	0.817	0.536	0.808	0.717	0.819	0.836

<b>kNN</b>	0.618	0.711	0.589	0.686	0.622	0.712	0.745
<b>Naïve bayes</b>	0.793	0.684	0.615	0.714	0.75	0.684	0.714
<b>Random forest</b>	0.642	0.762	0.617	0.747	0.667	0.755	0.774
<b>SGDClassifier_L1</b>	0.709	0.692	0.521	0.685	0.694	0.696	0.697
<b>SGDClassifier_L2</b>	0.709	0.692	0.521	0.711	0.608	0.696	0.697
<b>SGDClassifier_EN</b>	0.577	0.727	0.521	0.685	0.679	0.707	0.697
<b>LinearSVC_L1</b>	0.684	0.722	0.624	0.725	0.688	0.734	0.737
<b>LinearSVC_L2</b>	0.681	0.722	0.612	0.721	0.688	0.734	0.733

AdaBoost: Adaptive Boosting; Bagging: Bootstrap aggregating; SVC\_rbf or \_poly or \_sigmoid: support vector machine with rbf kernel or polynomial kernel or sigmoid kernel;

SGDClassifier\_L1 or \_L2 or \_EN: stochastic gradient descent with L1 or L2 or Elastic Net regularization; NN: k-Nearest Neighbors; LinearSVC\_L1 or \_L2: support vector machine with linear kernel coupled with L1 or L2 regularization

Sci: features extracted by scispaCy; Meta: features extracted by MetaMap; Custom: features extracted by the spaCy customized model; Combine1: features extracted by scispaCy or MetaMap; Combine2: features extracted by scispaCy or the spaCy customized model; Combine3: features extracted by MetaMap or the spaCy customized model; All: features extracted by scispaCy, MetaMap, or the spaCy customized model.

**Table S11: Precisions of the 15 machine learning configurations on the testing set.**

	SciSpaCy	MetaMap	Custom	Combine1	Combine2	Combine3	All
<b>Logistic regression</b>	0.667	0.757	0.599	0.767	0.642	0.78	0.777
<b>AdaBoost</b>	0.651	0.793	0.626	0.837	0.704	0.807	0.837
<b>Bagging</b>	0.658	0.757	0.634	0.761	0.667	0.771	0.779
<b>Decision tree</b>	0.633	0.703	0.613	0.679	0.603	0.711	0.695

<b>SVC_rbf</b>	0.636	0.792	0.614	0.792	0.634	0.773	0.778
<b>SVC_poly</b>	0.572	0.585	0.541	0.578	0.571	0.585	0.578
<b>SVC_sigmoid</b>	0.589	0.649	0.573	0.662	0.588	0.667	0.669
<b>kNN</b>	0.642	0.732	0.61	0.748	0.679	0.722	0.721
<b>Naïve bayes</b>	0.566	0.572	0.586	0.578	0.567	0.571	0.578
<b>Random forest</b>	0.719	0.832	0.63	0.812	0.733	0.818	0.829
<b>SGDClassifier_L1</b>	0.574	0.714	0.564	0.723	0.592	0.748	0.709
<b>SGDClassifier_L2</b>	0.569	0.788	0.503	0.745	0.578	0.787	0.786
<b>SGDClassifier_EN</b>	0.560	0.753	0.568	0.736	0.568	0.728	0.738
<b>LinearSVC_L1</b>	0.657	0.764	0.595	0.748	0.648	0.757	0.74
<b>LinearSVC_L2</b>	0.657	0.764	0.595	0.75	0.648	0.758	0.745

AdaBoost: Adaptive Boosting; Bagging: Bootstrap aggregating; SVC\_rbf or \_poly or \_sigmoid: support vector machine with rbf kernel or polynomial kernel or sigmoid kernel; SGDClassifier\_L1 or \_L2 or \_EN: stochastic gradient descent with L1 or L2 or Elastic Net regularization; NN: k-Nearest Neighbors; LinearSVC\_L1 or \_L2: support vector machine with linear kernel coupled with L1 or L2 regularization

Sci: features extracted by scispaCy; Meta: features extracted by MetaMap; Custom: features extracted by the spaCy customized model; Combine1: features extracted by scispaCy or MetaMap; Combine2: features extracted by scispaCy or the spaCy customized model; Combine3: features extracted by MetaMap or the spaCy customized model; All: features extracted by scispaCy, MetaMap, or the spaCy customized model.