# supp

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| before | default | tfidf | Logistic Regression | 0.9601226993865031 | 0.9567567567567568 | 0.4392059553349876 | 0.6020408163265305 | 0.9758939699735969 | before + default + tfidf |
| before | default | tfidf | K-Nearest Neighbors | 0.9335378323108384 | 0.8421052631578947 | 0.03970223325062035 | 0.07582938388625592 | 0.5846680545497062 | before + default + tfidf |
| before | default | tfidf | Support Vector Machine | 0.8203817314246762 | 0.24228028503562946 | 0.7593052109181141 | 0.36734693877551017 | 0.8621947924872695 | before + default + tfidf |
| before | default | tfidf | Naive Bayes | 0.96813224267212 | 0.8941605839416058 | 0.607940446650124 | 0.723781388478582 | 0.9738716261161144 | before + default + tfidf |
| before | default | tfidf | Decision Tree | 0.9444444444444444 | 0.9529411764705882 | 0.20099255583126552 | 0.3319672131147541 | 0.6471795477196416 | before + default + tfidf |
| before | default | tfidf | Random Forest | 0.9367757327880027 | 1.0 | 0.0794044665012407 | 0.1471264367816092 | 0.9178971074670985 | before + default + tfidf |
| before | default | tfidf\_pos | Logistic Regression | 0.9597818677573279 | 0.9417989417989417 | 0.44168734491315137 | 0.6013513513513513 | 0.9758535594205399 | before + default + tfidf + pos |
| before | default | tfidf\_pos | K-Nearest Neighbors | 0.935412406271302 | 0.8157894736842105 | 0.07692307692307693 | 0.1405895691609977 | 0.6104876736461897 | before + default + tfidf + pos |
| before | default | tfidf\_pos | Support Vector Machine | 0.8411724608043627 | 0.25117591721542804 | 0.6625310173697271 | 0.364256480218281 | 0.8263336050072763 | before + default + tfidf + pos |
| before | default | tfidf\_pos | Naive Bayes | 0.9625085207907293 | 0.8060200668896321 | 0.598014888337469 | 0.6866096866096866 | 0.9686677457949187 | before + default + tfidf + pos |
| before | default | tfidf\_pos | Decision Tree | 0.9444444444444444 | 0.8969072164948454 | 0.21588089330024815 | 0.34800000000000003 | 0.6394658996229106 | before + default + tfidf + pos |
| before | default | tfidf\_pos | Random Forest | 0.9369461486025903 | 1.0 | 0.08188585607940446 | 0.1513761467889908 | 0.9150769957251085 | before + default + tfidf + pos |
| before | default | tfidf\_pos\_liwc | Logistic Regression | 0.940865712338105 | 0.6647058823529411 | 0.2803970223325062 | 0.39441535776614306 | 0.9336043261994329 | before + default + tfidf + pos + liwc |
| before | default | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9350715746421268 | 0.7037037037037037 | 0.09429280397022333 | 0.1663019693654267 | 0.7010375068959019 | before + default + tfidf + pos + liwc |
| before | default | tfidf\_pos\_liwc | Support Vector Machine | 0.10650988411724609 | 0.06412103746397695 | 0.8833746898263027 | 0.11956339210747272 | 0.5356868318353429 | before + default + tfidf + pos + liwc |
| before | default | tfidf\_pos\_liwc | Naive Bayes | 0.9543285616905249 | 0.7303754266211604 | 0.5310173697270472 | 0.6149425287356322 | 0.9599102340860746 | before + default + tfidf + pos + liwc |
| before | default | tfidf\_pos\_liwc | Decision Tree | 0.9444444444444444 | 0.8811881188118812 | 0.22084367245657568 | 0.35317460317460314 | 0.6274360412187641 | before + default + tfidf + pos + liwc |
| before | default | tfidf\_pos\_liwc | Random Forest | 0.9364349011588275 | 1.0 | 0.07444168734491315 | 0.13856812933025403 | 0.9140283645758368 | before + default + tfidf + pos + liwc |
| before | under | tfidf | Logistic Regression | 0.9865410497981157 | 0.9787798408488063 | 0.9946091644204852 | 0.9866310160427807 | 0.9978769962032287 | before + under + tfidf |
| before | under | tfidf | K-Nearest Neighbors | 0.5423956931359354 | 1.0 | 0.08355795148247978 | 0.15422885572139303 | 0.9681404515549373 | before + under + tfidf |
| before | under | tfidf | Support Vector Machine | 0.892328398384926 | 0.8269662921348314 | 0.9919137466307277 | 0.9019607843137255 | 0.9847621945917746 | before + under + tfidf |
| before | under | tfidf | Naive Bayes | 0.9798115746971736 | 0.9863387978142076 | 0.9730458221024259 | 0.9796472184531886 | 0.9989131379880011 | before + under + tfidf |
| before | under | tfidf | Decision Tree | 0.6810228802153432 | 0.6109271523178808 | 0.9946091644204852 | 0.7569230769230769 | 0.7368996898820391 | before + under + tfidf |
| before | under | tfidf | Random Forest | 0.9098250336473755 | 0.9662576687116564 | 0.8490566037735849 | 0.9038737446197992 | 0.9823240008115236 | before + under + tfidf |
| before | under | tfidf\_pos | Logistic Regression | 0.9865410497981157 | 0.9787798408488063 | 0.9946091644204852 | 0.9866310160427807 | 0.9979422079239486 | before + under + tfidf + pos |
| before | under | tfidf\_pos | K-Nearest Neighbors | 0.5921938088829072 | 1.0 | 0.18328840970350405 | 0.30979498861047833 | 0.9698214648001623 | before + under + tfidf + pos |
| before | under | tfidf\_pos | Support Vector Machine | 0.9246298788694481 | 0.8776978417266187 | 0.9865229110512129 | 0.9289340101522843 | 0.9807118221603919 | before + under + tfidf + pos |
| before | under | tfidf\_pos | Naive Bayes | 0.9771197846567967 | 0.9862637362637363 | 0.967654986522911 | 0.9768707482993197 | 0.9988841550010144 | before + under + tfidf + pos |
| before | under | tfidf\_pos | Decision Tree | 0.6756393001345895 | 0.6069078947368421 | 0.9946091644204852 | 0.7538304392236977 | 0.7235783844883054 | before + under + tfidf + pos |
| before | under | tfidf\_pos | Random Forest | 0.9192462987886945 | 0.9905362776025236 | 0.8463611859838275 | 0.9127906976744186 | 0.9876496246703185 | before + under + tfidf + pos |
| before | under | tfidf\_pos\_liwc | Logistic Regression | 0.9744279946164199 | 0.9888888888888889 | 0.9595687331536388 | 0.9740082079343365 | 0.9961814914645103 | before + under + tfidf + pos + liwc |
| before | under | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.7523553162853297 | 0.8006430868167203 | 0.6711590296495957 | 0.7302052785923755 | 0.8147552386748979 | before + under + tfidf + pos + liwc |
| before | under | tfidf\_pos\_liwc | Support Vector Machine | 0.47510094212651416 | 0.4866385372714487 | 0.9326145552560647 | 0.6395563770794824 | 0.5880611830855288 | before + under + tfidf + pos + liwc |
| before | under | tfidf\_pos\_liwc | Naive Bayes | 0.8061911170928667 | 0.7734939759036145 | 0.8652291105121294 | 0.8167938931297709 | 0.8925890502275164 | before + under + tfidf + pos + liwc |
| before | under | tfidf\_pos\_liwc | Decision Tree | 0.6796769851951547 | 0.6113902847571189 | 0.9838274932614556 | 0.7541322314049588 | 0.735457786279454 | before + under + tfidf + pos + liwc |
| before | under | tfidf\_pos\_liwc | Random Forest | 0.8896366083445492 | 0.976897689768977 | 0.7978436657681941 | 0.8783382789317509 | 0.9824942758600701 | before + under + tfidf + pos + liwc |
| before | over | tfidf | Logistic Regression | 0.9780138339920948 | 0.974499803844645 | 0.9564882556796304 | 0.9654100272055965 | 0.9958092285900968 | before + over + tfidf |
| before | over | tfidf | K-Nearest Neighbors | 0.9961709486166008 | 0.9919478527607362 | 0.9961494031574895 | 0.9940441882804996 | 0.9987052639458442 | before + over + tfidf |
| before | over | tfidf | Support Vector Machine | 0.7225790513833992 | 0.7057444314185228 | 0.23180592991913745 | 0.3489855072463768 | 0.7295568074371768 | before + over + tfidf |
| before | over | tfidf | Naive Bayes | 0.9576333992094862 | 0.9472222222222222 | 0.9191374663072777 | 0.9329685362517101 | 0.9869486544373279 | before + over + tfidf |
| before | over | tfidf | Decision Tree | 0.7540760869565217 | 0.9220055710306406 | 0.254909510974201 | 0.3993966817496229 | 0.6231722531831495 | before + over + tfidf |
| before | over | tfidf | Random Forest | 0.7537055335968379 | 0.9886547811993517 | 0.23488640739314592 | 0.37958929682638454 | 0.930330210911733 | before + over + tfidf |
| before | over | tfidf\_pos | Logistic Regression | 0.9782608695652174 | 0.9748921145547274 | 0.9568733153638814 | 0.965798678585309 | 0.9955834025341396 | before + over + tfidf + pos |
| before | over | tfidf\_pos | K-Nearest Neighbors | 0.9930830039525692 | 0.982529434105583 | 0.9961494031574895 | 0.9892925430210324 | 0.996174121482374 | before + over + tfidf + pos |
| before | over | tfidf\_pos | Support Vector Machine | 0.7024456521739131 | 0.7126696832579186 | 0.12129380053908356 | 0.2073050345508391 | 0.7257952105689675 | before + over + tfidf + pos |
| before | over | tfidf\_pos | Naive Bayes | 0.9544219367588933 | 0.9431185361972951 | 0.9129765113592607 | 0.9278027783212679 | 0.9833430701125832 | before + over + tfidf + pos |
| before | over | tfidf\_pos | Decision Tree | 0.7591403162055336 | 0.8345398138572906 | 0.3107431651906045 | 0.4528619528619528 | 0.6438097086717836 | before + over + tfidf + pos |
| before | over | tfidf\_pos | Random Forest | 0.7553112648221344 | 0.9873417721518988 | 0.24027724297266076 | 0.38649736760606995 | 0.9349706737732201 | before + over + tfidf + pos |
| before | over | tfidf\_pos\_liwc | Logistic Regression | 0.915143280632411 | 0.9400921658986175 | 0.7855217558721602 | 0.8558842039018251 | 0.9725132227282827 | before + over + tfidf + pos + liwc |
| before | over | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9204545454545454 | 0.8029785913744958 | 0.9965344628417405 | 0.8893470790378007 | 0.9970216519221509 | before + over + tfidf + pos + liwc |
| before | over | tfidf\_pos\_liwc | Support Vector Machine | 0.3155879446640316 | 0.3064176749079432 | 0.8971890643049673 | 0.45681795902362515 | 0.5342369106491375 | before + over + tfidf + pos + liwc |
| before | over | tfidf\_pos\_liwc | Naive Bayes | 0.9308300395256917 | 0.9259723964868256 | 0.8525221409318444 | 0.8877305533279871 | 0.9734168070464452 | before + over + tfidf + pos + liwc |
| before | over | tfidf\_pos\_liwc | Decision Tree | 0.7680335968379447 | 0.8878101402373247 | 0.31690412013862146 | 0.46708286038592506 | 0.6519439982191602 | before + over + tfidf + pos + liwc |
| before | over | tfidf\_pos\_liwc | Random Forest | 0.7540760869565217 | 0.9918831168831169 | 0.235271467077397 | 0.3803299097416745 | 0.925181516883071 | before + over + tfidf + pos + liwc |
| simple | default | tfidf | Logistic Regression | 0.965576005453306 | 0.9276595744680851 | 0.5409429280397022 | 0.6833855799373041 | 0.9786073796934701 | simple + default + tfidf |
| simple | default | tfidf | K-Nearest Neighbors | 0.9350715746421268 | 0.9583333333333334 | 0.05707196029776675 | 0.10772833723653395 | 0.6181386626831246 | simple + default + tfidf |
| simple | default | tfidf | Support Vector Machine | 0.8738922972051807 | 0.3285859613428281 | 0.8014888337468983 | 0.46608946608946616 | 0.9013928019269931 | simple + default + tfidf |
| simple | default | tfidf | Naive Bayes | 0.9700068166325835 | 0.8900343642611683 | 0.6426799007444168 | 0.7463976945244957 | 0.975342297816695 | simple + default + tfidf |
| simple | default | tfidf | Decision Tree | 0.9473415132924335 | 0.9051724137931034 | 0.26054590570719605 | 0.4046242774566474 | 0.6447937359102249 | simple + default + tfidf |
| simple | default | tfidf | Random Forest | 0.9395023858214042 | 1.0 | 0.11910669975186104 | 0.21286031042128603 | 0.9471275588620571 | simple + default + tfidf |
| simple | default | tfidf\_pos | Logistic Regression | 0.9657464212678937 | 0.9208333333333333 | 0.5483870967741935 | 0.6874027993779159 | 0.9783876189330252 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos | K-Nearest Neighbors | 0.9352419904567144 | 0.8285714285714286 | 0.07196029776674938 | 0.1324200913242009 | 0.612348602317023 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos | Support Vector Machine | 0.7397750511247444 | 0.18355855855855857 | 0.8089330024813896 | 0.29921982560807714 | 0.8641508448756922 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos | Naive Bayes | 0.9630197682344922 | 0.790625 | 0.6277915632754343 | 0.6998616874135546 | 0.9684071204302588 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos | Decision Tree | 0.943762781186094 | 0.8924731182795699 | 0.20595533498759305 | 0.3346774193548387 | 0.6738539181209546 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos | Random Forest | 0.9395023858214042 | 1.0 | 0.11910669975186104 | 0.21286031042128603 | 0.9464206012091383 | simple + default + tfidf + pos |
| simple | default | tfidf\_pos\_liwc | Logistic Regression | 0.9567143830947512 | 0.8281938325991189 | 0.4665012406947891 | 0.5968253968253969 | 0.9659820331956801 | simple + default + tfidf + pos + liwc |
| simple | default | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9364349011588275 | 0.6875 | 0.13647642679900746 | 0.2277432712215321 | 0.8087541063251595 | simple + default + tfidf + pos + liwc |
| simple | default | tfidf\_pos\_liwc | Support Vector Machine | 0.1252556237218814 | 0.06252312245652979 | 0.8387096774193549 | 0.11637114821828198 | 0.3570435821003953 | simple + default + tfidf + pos + liwc |
| simple | default | tfidf\_pos\_liwc | Naive Bayes | 0.940013633265167 | 0.571830985915493 | 0.5037220843672456 | 0.5356200527704486 | 0.9554164897759031 | simple + default + tfidf + pos + liwc |
| simple | default | tfidf\_pos\_liwc | Decision Tree | 0.9452965235173824 | 0.8306451612903226 | 0.2555831265508685 | 0.39089184060721066 | 0.6649286345092502 | simple + default + tfidf + pos + liwc |
| simple | default | tfidf\_pos\_liwc | Random Forest | 0.939161554192229 | 1.0 | 0.1141439205955335 | 0.20489977728285078 | 0.9459059342216087 | simple + default + tfidf + pos + liwc |
| simple | under | tfidf | Logistic Regression | 0.9757738896366084 | 0.9862258953168044 | 0.9649595687331537 | 0.9754768392370572 | 0.9981342202127351 | simple + under + tfidf |
| simple | under | tfidf | K-Nearest Neighbors | 0.5195154777927322 | 1.0 | 0.03773584905660377 | 0.07272727272727272 | 0.9781069762049678 | simple + under + tfidf |
| simple | under | tfidf | Support Vector Machine | 0.8855989232839838 | 0.8191964285714286 | 0.9892183288409704 | 0.8962148962148963 | 0.9829471350317365 | simple + under + tfidf |
| simple | under | tfidf | Naive Bayes | 0.9744279946164199 | 0.9861878453038674 | 0.9622641509433962 | 0.9740791268758526 | 0.9982066776802017 | simple + under + tfidf |
| simple | under | tfidf | Decision Tree | 0.7281292059219381 | 0.9941520467836257 | 0.4582210242587601 | 0.6273062730627306 | 0.7630097382836275 | simple + under + tfidf |
| simple | under | tfidf | Random Forest | 0.917900403768506 | 0.9936305732484076 | 0.8409703504043127 | 0.9109489051094891 | 0.9881858099295713 | simple + under + tfidf |
| simple | under | tfidf\_pos | Logistic Regression | 0.9757738896366084 | 0.9862258953168044 | 0.9649595687331537 | 0.9754768392370572 | 0.998177694693215 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos | K-Nearest Neighbors | 0.5612382234185733 | 1.0 | 0.12129380053908356 | 0.21634615384615385 | 0.9780852389647277 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos | Support Vector Machine | 0.917900403768506 | 0.8780487804878049 | 0.9703504043126685 | 0.9218950064020488 | 0.9841861577254152 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos | Naive Bayes | 0.971736204576043 | 0.9915730337078652 | 0.9514824797843666 | 0.9711141678129298 | 0.9980472712517753 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos | Decision Tree | 0.7281292059219381 | 0.9828571428571429 | 0.4636118598382749 | 0.6300366300366301 | 0.7640821088021331 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos | Random Forest | 0.8990578734858681 | 0.9933333333333333 | 0.8032345013477089 | 0.8882265275707899 | 0.986819986667826 | simple + under + tfidf + pos |
| simple | under | tfidf\_pos\_liwc | Logistic Regression | 0.9730820995962315 | 0.9943661971830986 | 0.9514824797843666 | 0.9724517906336088 | 0.9980364026316553 | simple + under + tfidf + pos + liwc |
| simple | under | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.8317631224764468 | 0.875 | 0.7735849056603774 | 0.821173104434907 | 0.914000231863896 | simple + under + tfidf + pos + liwc |
| simple | under | tfidf\_pos\_liwc | Support Vector Machine | 0.4253028263795424 | 0.45894428152492667 | 0.8436657681940701 | 0.5944919278252612 | 0.47677013593020895 | simple + under + tfidf + pos + liwc |
| simple | under | tfidf\_pos\_liwc | Naive Bayes | 0.8613728129205922 | 0.932258064516129 | 0.7789757412398922 | 0.8487518355359766 | 0.9520331565371126 | simple + under + tfidf + pos + liwc |
| simple | under | tfidf\_pos\_liwc | Decision Tree | 0.7052489905787349 | 1.0 | 0.40970350404312667 | 0.5812619502868068 | 0.7415876880271279 | simple + under + tfidf + pos + liwc |
| simple | under | tfidf\_pos\_liwc | Random Forest | 0.9057873485868102 | 0.9966996699669967 | 0.8140161725067385 | 0.8961424332344212 | 0.9898161029475697 | simple + under + tfidf + pos + liwc |
| simple | over | tfidf | Logistic Regression | 0.9740612648221344 | 0.9693275658670861 | 0.9491721216788602 | 0.9591439688715954 | 0.9945520251765592 | simple + over + tfidf |
| simple | over | tfidf | K-Nearest Neighbors | 0.9940711462450593 | 0.9847850893875999 | 0.9969195225259915 | 0.9908151549942594 | 0.9995462821923795 | simple + over + tfidf |
| simple | over | tfidf | Support Vector Machine | 0.7982954545454546 | 0.8428165007112376 | 0.45629572583750483 | 0.5920559580314764 | 0.8647531252050379 | simple + over + tfidf |
| simple | over | tfidf | Naive Bayes | 0.9496047430830039 | 0.936229573535273 | 0.9045051983057374 | 0.9200940070505287 | 0.9839648795317776 | simple + over + tfidf |
| simple | over | tfidf | Decision Tree | 0.7949604743083004 | 0.8940285954583683 | 0.4093184443588756 | 0.5615425250924457 | 0.6976139393986501 | simple + over + tfidf |
| simple | over | tfidf | Random Forest | 0.8052124505928854 | 0.980225988700565 | 0.4008471313053523 | 0.5690079256627494 | 0.9514781383222055 | simple + over + tfidf |
| simple | over | tfidf\_pos | Logistic Regression | 0.9728260869565217 | 0.9680976762504924 | 0.9464767038891028 | 0.9571651090342679 | 0.994431094448299 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos | K-Nearest Neighbors | 0.9904891304347826 | 0.9740406320541761 | 0.9969195225259915 | 0.9853472882968601 | 0.9990708570739539 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos | Support Vector Machine | 0.7164031620553359 | 0.8028169014084507 | 0.15363881401617252 | 0.2579185520361991 | 0.781309732304743 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos | Naive Bayes | 0.9456521739130435 | 0.933306548814785 | 0.8944936465152099 | 0.9134880062917813 | 0.9795695692352226 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos | Decision Tree | 0.7796442687747036 | 0.9012833168805529 | 0.3515594917212168 | 0.5058171745152354 | 0.670123625935979 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos | Random Forest | 0.7945899209486166 | 0.9844398340248963 | 0.3654216403542549 | 0.532996349340073 | 0.950319563125665 | simple + over + tfidf + pos |
| simple | over | tfidf\_pos\_liwc | Logistic Regression | 0.9452816205533597 | 0.949874686716792 | 0.875625721986908 | 0.9112402324183531 | 0.9842836268826977 | simple + over + tfidf + pos + liwc |
| simple | over | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9423171936758893 | 0.848951507208388 | 0.9976896418944936 | 0.9173305009736236 | 0.9977032264696427 | simple + over + tfidf + pos + liwc |
| simple | over | tfidf\_pos\_liwc | Support Vector Machine | 0.5233448616600791 | 0.3684868695289704 | 0.6807855217558721 | 0.4781609195402299 | 0.5860128032520072 | simple + over + tfidf + pos + liwc |
| simple | over | tfidf\_pos\_liwc | Naive Bayes | 0.9055088932806324 | 0.8989547038327527 | 0.7947631882941856 | 0.8436541998773759 | 0.9624467024249097 | simple + over + tfidf + pos + liwc |
| simple | over | tfidf\_pos\_liwc | Decision Tree | 0.7787796442687747 | 0.9467849223946785 | 0.3288409703504043 | 0.48813946841954853 | 0.6612009758766655 | simple + over + tfidf + pos + liwc |
| simple | over | tfidf\_pos\_liwc | Random Forest | 0.8108942687747036 | 0.9872029250457038 | 0.4158644589911436 | 0.5852072609049038 | 0.9450724859625473 | simple + over + tfidf + pos + liwc |
| large | default | tfidf | Logistic Regression | 0.9645535105657805 | 0.9372197309417041 | 0.5186104218362283 | 0.6677316293929714 | 0.9840394207215327 | large + default + tfidf |
| large | default | tfidf | K-Nearest Neighbors | 0.9342194955691888 | 1.0 | 0.04218362282878412 | 0.08095238095238096 | 0.6205471770504384 | large + default + tfidf |
| large | default | tfidf | Support Vector Machine | 0.880879345603272 | 0.33555555555555555 | 0.749379652605459 | 0.46354566385264767 | 0.9009026537019926 | large + default + tfidf |
| large | default | tfidf | Naive Bayes | 0.971881390593047 | 0.8814102564102564 | 0.6823821339950372 | 0.7692307692307692 | 0.9824965548868391 | large + default + tfidf |
| large | default | tfidf | Decision Tree | 0.950068166325835 | 0.8666666666666667 | 0.3225806451612903 | 0.47016274864376134 | 0.6599181799813385 | large + default + tfidf |
| large | default | tfidf | Random Forest | 0.9398432174505794 | 1.0 | 0.12406947890818859 | 0.22075055187637968 | 0.940765621062525 | large + default + tfidf |
| large | default | tfidf\_pos | Logistic Regression | 0.9643830947511929 | 0.9217391304347826 | 0.5260545905707196 | 0.669826224328594 | 0.9836014883796957 | large + default + tfidf + pos |
| large | default | tfidf\_pos | K-Nearest Neighbors | 0.9357532379004772 | 0.8421052631578947 | 0.0794044665012407 | 0.14512471655328799 | 0.6122639217760665 | large + default + tfidf + pos |
| large | default | tfidf\_pos | Support Vector Machine | 0.880879345603272 | 0.3310502283105023 | 0.7196029776674938 | 0.4534792806880376 | 0.8888139956728924 | large + default + tfidf + pos |
| large | default | tfidf\_pos | Naive Bayes | 0.9657464212678937 | 0.7970588235294118 | 0.6724565756823822 | 0.7294751009421265 | 0.9766772082210502 | large + default + tfidf + pos |
| large | default | tfidf\_pos | Decision Tree | 0.9485344239945467 | 0.8914728682170543 | 0.2853598014888337 | 0.4323308270676692 | 0.6656848567128059 | large + default + tfidf + pos |
| large | default | tfidf\_pos | Random Forest | 0.9398432174505794 | 1.0 | 0.12406947890818859 | 0.22075055187637968 | 0.941973851193814 | large + default + tfidf + pos |
| large | default | tfidf\_pos\_liwc | Logistic Regression | 0.9531356509884117 | 0.8106796116504854 | 0.4143920595533499 | 0.548440065681445 | 0.9659702278655734 | large + default + tfidf + pos + liwc |
| large | default | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9377982276755283 | 0.7261904761904762 | 0.1513647642679901 | 0.2505133470225873 | 0.8203591998710494 | large + default + tfidf + pos + liwc |
| large | default | tfidf\_pos\_liwc | Support Vector Machine | 0.12951601908657123 | 0.06281360341943877 | 0.8387096774193549 | 0.11687413554633473 | 0.5082603256908956 | large + default + tfidf + pos + liwc |
| large | default | tfidf\_pos\_liwc | Naive Bayes | 0.9476823449216087 | 0.6379310344827587 | 0.5508684863523573 | 0.5912117177097205 | 0.9672483818751858 | large + default + tfidf + pos + liwc |
| large | default | tfidf\_pos\_liwc | Decision Tree | 0.9478527607361963 | 0.825503355704698 | 0.3052109181141439 | 0.44565217391304346 | 0.6515886114888564 | large + default + tfidf + pos + liwc |
| large | default | tfidf\_pos\_liwc | Random Forest | 0.939161554192229 | 1.0 | 0.1141439205955335 | 0.20489977728285078 | 0.9374782906790108 | large + default + tfidf + pos + liwc |
| large | under | tfidf | Logistic Regression | 0.9771197846567967 | 0.9836065573770492 | 0.9703504043126685 | 0.9769335142469471 | 0.997594412080109 | large + under + tfidf |
| large | under | tfidf | K-Nearest Neighbors | 0.5450874831763123 | 1.0 | 0.0889487870619946 | 0.16336633663366337 | 0.9849216010202011 | large + under + tfidf |
| large | under | tfidf | Support Vector Machine | 0.9286675639300135 | 0.8916256157635468 | 0.9757412398921833 | 0.9317889317889317 | 0.9855302437469206 | large + under + tfidf |
| large | under | tfidf | Naive Bayes | 0.9757738896366084 | 0.9943977591036415 | 0.9568733153638814 | 0.9752747252747253 | 0.9980508941251486 | large + under + tfidf |
| large | under | tfidf | Decision Tree | 0.6742934051144011 | 0.6055646481178396 | 0.9973045822102425 | 0.7535641547861508 | 0.7147929165579805 | large + under + tfidf |
| large | under | tfidf | Random Forest | 0.8963660834454913 | 0.9932885906040269 | 0.7978436657681941 | 0.8849028400597908 | 0.9851788250297075 | large + under + tfidf |
| large | under | tfidf\_pos | Logistic Regression | 0.9784656796769852 | 0.9836512261580381 | 0.9730458221024259 | 0.978319783197832 | 0.9976886067878156 | large + under + tfidf + pos |
| large | under | tfidf\_pos | K-Nearest Neighbors | 0.5841184387617766 | 1.0 | 0.16711590296495957 | 0.2863741339491917 | 0.9852259223835608 | large + under + tfidf + pos |
| large | under | tfidf\_pos | Support Vector Machine | 0.9138627187079408 | 0.8698795180722891 | 0.9730458221024259 | 0.9185750636132315 | 0.9840593571573485 | large + under + tfidf + pos |
| large | under | tfidf\_pos | Naive Bayes | 0.9771197846567967 | 0.9971910112359551 | 0.9568733153638814 | 0.9766162310866575 | 0.9979784366576819 | large + under + tfidf + pos |
| large | under | tfidf\_pos | Decision Tree | 0.6567967698519516 | 1.0 | 0.31266846361185985 | 0.4763860369609856 | 0.7306357418195519 | large + under + tfidf + pos |
| large | under | tfidf\_pos | Random Forest | 0.8815612382234186 | 0.9829351535836177 | 0.7762803234501348 | 0.8674698795180723 | 0.9818856331333506 | large + under + tfidf + pos |
| large | under | tfidf\_pos\_liwc | Logistic Regression | 0.9784656796769852 | 0.9944289693593314 | 0.9622641509433962 | 0.9780821917808219 | 0.9978987334434687 | large + under + tfidf + pos + liwc |
| large | under | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.8209959623149394 | 0.8695652173913043 | 0.7547169811320755 | 0.8080808080808081 | 0.915119699736255 | large + under + tfidf + pos + liwc |
| large | under | tfidf\_pos\_liwc | Support Vector Machine | 0.4253028263795424 | 0.45894428152492667 | 0.8436657681940701 | 0.5944919278252612 | 0.48364634959278907 | large + under + tfidf + pos + liwc |
| large | under | tfidf\_pos\_liwc | Naive Bayes | 0.8532974427994616 | 0.9119496855345912 | 0.7816711590296496 | 0.8417997097242379 | 0.9523881981276989 | large + under + tfidf + pos + liwc |
| large | under | tfidf\_pos\_liwc | Decision Tree | 0.648721399730821 | 1.0 | 0.29649595687331537 | 0.45738045738045735 | 0.7168724458742717 | large + under + tfidf + pos + liwc |
| large | under | tfidf\_pos\_liwc | Random Forest | 0.8842530282637954 | 0.9830508474576272 | 0.7816711590296496 | 0.8708708708708709 | 0.9829435121583631 | large + under + tfidf + pos + liwc |
| large | over | tfidf | Logistic Regression | 0.9745553359683794 | 0.9730906212900673 | 0.9468617635733538 | 0.9597970335675254 | 0.9957178128021736 | large + over + tfidf |
| large | over | tfidf | K-Nearest Neighbors | 0.9960474308300395 | 0.9904397705544933 | 0.9973045822102425 | 0.9938603223330775 | 0.9996590551731919 | large + over + tfidf |
| large | over | tfidf | Support Vector Machine | 0.7727272727272727 | 0.6029371770465053 | 0.8536773199845976 | 0.7067261715014344 | 0.88779081406827 | large + over + tfidf |
| large | over | tfidf | Naive Bayes | 0.9540513833992095 | 0.9392025266482432 | 0.9160569888332691 | 0.9274853801169591 | 0.9869056249454254 | large + over + tfidf |
| large | over | tfidf | Decision Tree | 0.7921195652173914 | 0.9279026217228464 | 0.3815941470927994 | 0.5407912687585266 | 0.6869384940153994 | large + over + tfidf |
| large | over | tfidf | Random Forest | 0.7978013833992095 | 0.9858299595141701 | 0.37504813246053137 | 0.5433751743375175 | 0.9506492341555711 | large + over + tfidf |
| large | over | tfidf\_pos | Logistic Regression | 0.9744318181818182 | 0.9715864246250987 | 0.948016942626107 | 0.9596569869421165 | 0.995592785694294 | large + over + tfidf + pos |
| large | over | tfidf\_pos | K-Nearest Neighbors | 0.9919713438735178 | 0.9780966767371602 | 0.9973045822102425 | 0.9876072449952336 | 0.999023801226015 | large + over + tfidf + pos |
| large | over | tfidf\_pos | Support Vector Machine | 0.7679100790513834 | 0.7643593519882179 | 0.3996919522525991 | 0.5249051833122629 | 0.8114499482280637 | large + over + tfidf + pos |
| large | over | tfidf\_pos | Naive Bayes | 0.9503458498023716 | 0.9356887653830885 | 0.9075856757797459 | 0.921422986708366 | 0.9829301060304101 | large + over + tfidf + pos |
| large | over | tfidf\_pos | Decision Tree | 0.7860671936758893 | 0.9538300104931794 | 0.35001925298421255 | 0.512112676056338 | 0.6725804033540456 | large + over + tfidf + pos |
| large | over | tfidf\_pos | Random Forest | 0.8002717391304348 | 0.9851485148514851 | 0.38313438582980364 | 0.5517050180205156 | 0.9519352522736133 | large + over + tfidf + pos |
| large | over | tfidf\_pos\_liwc | Logistic Regression | 0.9504693675889329 | 0.9590301003344481 | 0.8833269156719291 | 0.9196231709761474 | 0.9874393096851087 | large + over + tfidf + pos + liwc |
| large | over | tfidf\_pos\_liwc | K-Nearest Neighbors | 0.9400938735177866 | 0.8441981747066493 | 0.9973045822102425 | 0.9143865842894968 | 0.9981195866956032 | large + over + tfidf + pos + liwc |
| large | over | tfidf\_pos\_liwc | Support Vector Machine | 0.3158349802371542 | 0.3087623504940198 | 0.9145167500962649 | 0.4616580814462047 | 0.5613918811716598 | large + over + tfidf + pos + liwc |
| large | over | tfidf\_pos\_liwc | Naive Bayes | 0.9148962450592886 | 0.9045801526717557 | 0.8213323065075087 | 0.8609485368314833 | 0.9677047732905966 | large + over + tfidf + pos + liwc |
| large | over | tfidf\_pos\_liwc | Decision Tree | 0.7832262845849802 | 0.9488272921108742 | 0.3427031189834424 | 0.5035360678925036 | 0.6688438049050541 | large + over + tfidf + pos + liwc |
| large | over | tfidf\_pos\_liwc | Random Forest | 0.799407114624506 | 0.986013986013986 | 0.38005390835579517 | 0.5486381322957199 | 0.9479125374634924 | large + over + tfidf + pos + liwc |