Introduction:

Objectives:

Methods:

Workflow:

Features:

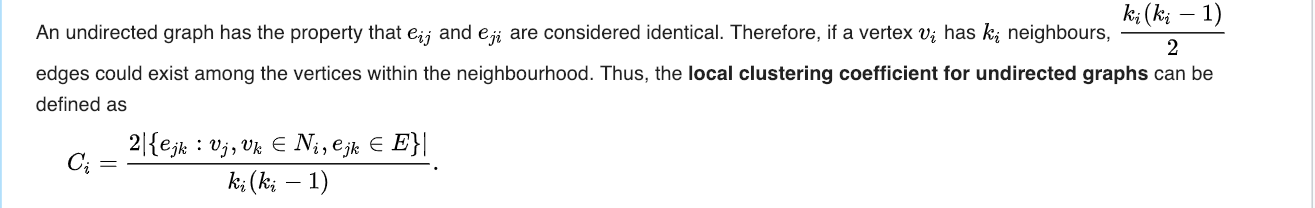
**Topological Features:**

1. Average Shortest Path:

Step 1: Find the geodesic distance from a non-disease node to each disease node

Step 2: Take the average of all distances as the Average Shortest Path of a non-disease node to all disease nodes

1. Local Clustering Coefficient:

Implemented according to the formula

The cluster E is the disease cluster, and N is the non-disease nodes cluster

1. Degree Centrality

Implemented by NetworkX library

Rationale: Higher, more likely to be involved in a more important functional module

1. Closeness Centrality

Implemented by NetworkX library

Rationale: Higher, more functionally important as needs to be communicated quickly.

1. Betweeness Centrality

Implemented by NetworkX library

Rationale: An important node will lie on a higher proportion of the paths.

1. Eigenvector Centrality

Implemented by NetworkX library

Rationale: the influence of a node in a network

1. Percolation centrality

Implemented by NetworkX library. This method has problems; thus the library method couldn’t successfully be run.

Rationale: importance of a node in purely topological terms, despite the network dynamic

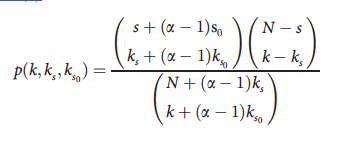
1. Pagerank

Implemented by NetworkX library.

Rationale: the notion of how central a node is in a network relative to a particular node

1. Connectivity Significance

Implemented using the formula given in the DIAMOnD paper



1. Articulation Point

Implemented using the algorithm given in this website (<https://www.geeksforgeeks.org/articulation-points-or-cut-vertices-in-a-graph/> ), the code has been slightly modified.

1. Modularity

Step 1: NetworkX library computes the modularity matrix

Step 2: Summing over all modularity from a node (row, in Gc.nodes()) to a disease node (column), exclude the node itself.

Step 3: Take the average for each node ( in Gc.nodes()).

**Sequence based features:**

**Functional Features:**

**Disease Module Identification**

**Dataset (After Data Cleaning, Data Balancing and Stratified Split)**

|  |  |  |
| --- | --- | --- |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 1441 | 4329 | 5770 |

**Dataset 1**

|  |  |  |
| --- | --- | --- |
| Train set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 1009 | 1008 | 2017 |
| Test set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 432 | 433 | 865 |

**Dataset 2**

|  |  |  |
| --- | --- | --- |
| Train set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 1009 | 1008 | 2017 |
| Test set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 432 | 433 | 865 |

**Dataset 3**

|  |  |  |
| --- | --- | --- |
| Train set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 1008 | 994 | 2002 |
| Test set | | |
| Disease Node | Non Disease Node (Essential and Non-essential) | Total |
| 433 | 426 | 859 |

**Experiment1: K means clustering with Principle Component Analysis (on the whole dataset)**

* Dataset 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Condition | Accuracy | Precision | Recall | F1 |
| Cluster 0 is considered as disease | 0.45107564 | 0.41198502 | 0.22900763 | 0.29438002 |
| Cluster 0 is considered as non-disease | 0.54892436 | 0.53387794 | 0.77099237 | 0.63089154 |
| Cluster 1 is considered as disease | 0.52602359 | 0.52856055 | 0.48160999 | 0.50399419 |
| Cluster 1 is considered as non-disease | 0.47397641 | 0.47609943 | 0.51839001 | 0.49634551 |
| Cluster 2 is considered as disease | 0.52290076 | 0.54296875 | 0.28938237 | 0.3775464 |
| Cluster 2 is considered as non-disease | 0.47709924 | 0.48438978 | 0.71061763 | 0.57609001 |

Conclusion: Disease set: cluster 1; Non-disease set: cluster 0 and cluster 2

* Dataset 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Condition | Accuracy | Precision | Recall | F1 |
| Cluster 0 is considered as disease | 0.45697432 | 0.42091837 | 0.22900763 | 0.29662921 |
| Cluster 0 is considered as non-disease | 0.54302568 | 0.52955195 | 0.77099237 | 0.62786098 |
| Cluster 1 is considered as disease | 0.51804303 | 0.53412073 | 0.28244275 | 0.36949614 |
| Cluster 1 is considered as non-disease | 0.48195697 | 0.48773585 | 0.71755725 | 0.58073575 |
| Cluster 2 is considered as disease | 0.52498265 | 0.52694611 | 0.48854962 | 0.50702197 |
| Cluster 2 is considered as non-disease | 0.47501735 | 0.4767141 | 0.51145038 | 0.49347171 |

Conclusion: Disease set: cluster 2; Non-disease set: cluster 0 and cluster 1

* Dataset 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Condition | Accuracy | Precision | Recall | F1 |
| Cluster 0 is considered as disease | 0.52079692 | 0.54475703 | 0.29562804 | 0.38326586 |
| Cluster 0 is considered as non-disease | 0.47920308 | 0.48821549 | 0.70437196 | 0.57670455 |
| Cluster 1 is considered as disease | 0.44669696 | 0.40482574 | 0.20957668 | 0.27617741 |
| Cluster 1 is considered as non-disease | 0.55330304 | 0.53853428 | 0.79042332 | 0.64060742 |
| Cluster 2 is considered as disease | 0.52883607 | 0.53488372 | 0.49479528 | 0.51405912 |
| Cluster 2 is considered as non-disease | 0.47116393 | 0.47643979 | 0.50520472 | 0.49040081 |

Conclusion: Disease set: cluster 2; Non-disease set: cluster 0 and cluster 1

General Conclusion for Kmeans:

Cluster 2 is the disease set, and cluster 0 and cluster 1 are non-disease cluster

**Experiment2: Support Vector Machine with RBF Kernel**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Accuracy | Precision | Recall | F1 Score |
| Dataset 1 | 0.55491329 | 0.55119826 | 0.58564815 | 0.56790123 |
| Dataset 2 | 0.59306358 | 0.5952381 | 0.5787037 | 0.58685446 |
| Dataset 3 | 0.56810244 | 0.562 | 0.64896074 | 0.60235798 |

* Dataset 1

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.62376238 |
| CV2 | 0.60891089 |
| CV3 | 0.61881188 |
| CV4 | 0.54455446 |
| CV5 | 0.58910891 |
| CV6 | 0.59405941 |
| CV7 | 0.58415842 |
| CV8 | 0.58415842 |
| CV9 | 0.51741294 |
| CV10 | 0.48 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.55491329 | 0.55119826 | 0.58564815 | 0.56790123 |

* Dataset 2

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.55445545 |
| CV2 | 0.53960396 |
| CV3 | 0.57920792 |
| CV4 | 0.56435644 |
| CV5 | 0.56930693 |
| CV6 | 0.55940594 |
| CV7 | 0.64356436 |
| CV8 | 0.55940594 |
| CV9 | 0.60696517 |
| CV10 | 0.55 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.59306358 | 0.5952381 | 0.5787037 | 0.58685446 |

* Dataset 3

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.51741294 |
| CV2 | 0.53233831 |
| CV3 | 0.59701493 |
| CV4 | 0.62686567 |
| CV5 | 0.66 |
| CV6 | 0.66 |
| CV7 | 0.61 |
| CV8 | 0.55 |
| CV9 | 0.6281407 |
| CV10 | 0.5678392 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.56810244 | 0.562 | 0.64896074 | 0.60235798 |

**Experiment3: Random Forest**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Accuracy | Precision | Recall | F1 Score |
| Dataset 1 | 0.68786127 | 0.68786127 | 0.68287037 | 0.68604651 |
| Dataset 2 | 0.69710983 | 0.69767442 | 0.69444444 | 0.69605568 |
| Dataset 3 | 0.67520373 | 0.6673913 | 0.70900693 | 0.68756999 |

* Dataset 1

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.71287129 |
| CV2 | 0.72277228 |
| CV3 | 0.70792079 |
| CV4 | 0.73267327 |
| CV5 | 0.68316832 |
| CV6 | 0.66831683 |
| CV7 | 0.72277228 |
| CV8 | 0.71287129 |
| CV9 | 0.68159204 |
| CV10 | 0.64 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.68786127 | 0.68786127 | 0.68287037 | 0.68604651 |

* Dataset 2

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.62871287 |
| CV2 | 0.62376238 |
| CV3 | 0.67821782 |
| CV4 | 0.65346535 |
| CV5 | 0.65841584 |
| CV6 | 0.62871287 |
| CV7 | 0.65841584 |
| CV8 | 0.65841584 |
| CV9 | 0.70646766 |
| CV10 | 0.66 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.69710983 | 0.69767442 | 0.69444444 | 0.69605568 |

* Dataset 3

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.6119403 |
| CV2 | 0.71641791 |
| CV3 | 0.69154229 |
| CV4 | 0.68159204 |
| CV5 | 0.715 |
| CV6 | 0.75 |
| CV7 | 0.7 |
| CV8 | 0.625 |
| CV9 | 0.69346734 |
| CV10 | 0.66834171 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.67520373 | 0.6673913 | 0.70900693 | 0.68756999 |

**Experiment4: Deep Learning**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Accuracy | Precision | Recall | F1 Score |
| Dataset 1 | 0.54219653 | 0.53501946 | 0.63657407 | 0.58139535 |
| Dataset 2 | 0.56878613 | 0.57040573 | 0.55324074 | 0.56169213 |
| Dataset 3 | 0.58207218 | 0.58447489 | 0.59122402 | 0.58783008 |

* Dataset 1

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.58910891 |
| CV2 | 0.58910891 |
| CV3 | 0.53465347 |
| CV4 | 0.52475248 |
| CV5 | 0.57920792 |
| CV6 | 0.56930693 |
| CV7 | 0.51485149 |
| CV8 | 0.59405941 |
| CV9 | 0.55223881 |
| CV10 | 0.52 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.54219653 | 0.53501946 | 0.63657407 | 0.58139535 |

* Dataset 2

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.56435644 |
| CV2 | 0.49009901 |
| CV3 | 0.54950495 |
| CV4 | 0.56435644 |
| CV5 | 0.61881188 |
| CV6 | 0.51485149 |
| CV7 | 0.53960396 |
| CV8 | 0.4950495 |
| CV9 | 0.60199005 |
| CV10 | 0.59 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.56878613 | 0.57040573 | 0.55324074 | 0.56169213 |

* **Dataset 3**

Cross Validation Score

|  |  |
| --- | --- |
| CV1 | 0.55223881 |
| CV2 | 0.48258706 |
| CV3 | 0.63681592 |
| CV4 | 0.56716418 |
| CV5 | 0.67 |
| CV6 | 0.605 |
| CV7 | 0.56 |
| CV8 | 0.555 |
| CV9 | 0.5678392 |
| CV10 | 0.5678392 |

Test Set Score

|  |  |  |  |
| --- | --- | --- | --- |
| Accuracy | Precision | Recall | F1 Score |
| 0.58207218 | 0.58447489 | 0.59122402 | 0.58783008 |

**Feature Selection**

**Feature Importance using the model that gives the best result**

Average is: 1/36 = 0.02777.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Dataset 1 | Dataset 2 | Dataset 3 | Average | Met  Average | Average Score Ranking |
| Average Shortest Path to all Disease genes | 0.04346083 | 0.03424906 | 0.031861 | 0.0365236 | 1 | 2 |
| BetweennessCentrality | 0.01667202 | 0.01703063 | 0.01930463 | 0.0176691 | 0 | 35 |
| ClosenessCentrality | 0.02228344 | 0.02340866 | 0.02744768 | 0.0243799 | 0 | 18 |
| DegreeCentrality | 0.01317933 | 0.01338562 | 0.01184739 | 0.0128041 | 0 | 36 |
| EigenvectorCentrality | 0.02404728 | 0.02516761 | 0.02079586 | 0.0233369 | 0 | 30 |
| HarmonicCentrality | 0.02559715 | 0.02451262 | 0.02254047 | 0.0242167 | 0 | 20 |
| Modularity | 0.12664342 | 0.11366804 | 0.13332279 | 0.1245448 | 1 | 1 |
| PageRank | 0.0245738 | 0.02959554 | 0.02306586 | 0.0257451 | 0 | 12 |
| BP | 0.02930925 | 0.04047054 | 0.03008127 | 0.033287 | 1 | 4 |
| CC | 0.02815931 | 0.02608422 | 0.02585718 | 0.0267002 | 0 | 7 |
| MF | 0.02505777 | 0.02997702 | 0.02251853 | 0.0258511 | 0 | 11 |
| FrequencyA | 0.02282336 | 0.02592521 | 0.02355146 | 0.0241 | 0 | 23 |
| FrequencyC | 0.02604764 | 0.0257333 | 0.02863298 | 0.0268046 | 0 | 6 |
| FrequencyD | 0.0230632 | 0.02453177 | 0.02503815 | 0.024211 | 0 | 21 |
| FrequencyE | 0.03329916 | 0.03024063 | 0.03877915 | 0.0341063 | 1 | 3 |
| FrequencyF | 0.02312303 | 0.02213389 | 0.0225686 | 0.0226085 | 0 | 33 |
| FrequencyG | 0.02781013 | 0.02360264 | 0.02279414 | 0.0247356 | 0 | 16 |
| FrequencyH | 0.02493357 | 0.02560892 | 0.02747443 | 0.0260056 | 0 | 10 |
| FrequencyI | 0.02552015 | 0.02288315 | 0.02295234 | 0.0237852 | 0 | 26 |
| FrequencyK | 0.02260602 | 0.0255323 | 0.02641962 | 0.0248526 | 0 | 15 |
| FrequencyL | 0.02359026 | 0.02509347 | 0.02257635 | 0.0237534 | 0 | 27 |
| FrequencyM | 0.02351522 | 0.02450154 | 0.02178546 | 0.0232674 | 0 | 31 |
| FrequencyN | 0.02481608 | 0.02290447 | 0.025167 | 0.0242959 | 0 | 19 |
| FrequencyP | 0.02556854 | 0.02590297 | 0.02413133 | 0.0252009 | 0 | 14 |
| FrequencyQ | 0.02476733 | 0.02491347 | 0.02291031 | 0.024197 | 0 | 22 |
| FrequencyR | 0.02260081 | 0.02390277 | 0.02531955 | 0.023941 | 0 | 25 |
| FrequencyS | 0.02403117 | 0.02507819 | 0.02660608 | 0.0252385 | 0 | 13 |
| FrequencyT | 0.02762817 | 0.02418846 | 0.02779702 | 0.0265379 | 0 | 8 |
| FrequencyV | 0.02291558 | 0.02496835 | 0.02426742 | 0.0240505 | 0 | 24 |
| FrequencyW | 0.02485183 | 0.02533806 | 0.02842926 | 0.0262064 | 0 | 9 |
| FrequencyX | 0.02477384 | 0.02698961 | 0.02235954 | 0.0247077 | 0 | 17 |
| Aromaticity | 0.0240424 | 0.02275794 | 0.02095999 | 0.0225868 | 0 | 34 |
| Isoelectric | 0.02889065 | 0.03005274 | 0.02974957 | 0.0295643 | 1 | 5 |
| SSfractionHelix | 0.02501649 | 0.02274011 | 0.02255569 | 0.0234374 | 0 | 29 |
| SSfractionTurn | 0.02226102 | 0.02332593 | 0.02377288 | 0.0231199 | 0 | 32 |
| SSfractionSheet | 0.02252077 | 0.02360054 | 0.02475901 | 0.0236268 | 0 | 28 |

**Independent T Test**

Threshold is 0.05, if p-value < 0.05, then it is normal.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Dataset 1 | Dataset 2 | Dataset 3 | Average/highest | Met |
| Average |
| Average Shortest Path to all Disease genes | 0.85498631 | 0.09209156 | 0.10143644 | 0.34950477 | 0 |
| BetweennessCentrality | 0.75285645 | 0.2008714 | 0.02114846 | 0.32495877 | 0 |
| ClosenessCentrality | 0.93733359 | 0.45606944 | 0.09490076 | 0.49610126 | 0 |
| DegreeCentrality | 0.76680924 | 0.98211418 | 0.97156584 | 0.90682975 | 0 |
| EigenvectorCentrality | 0.51673462 | 0.48240202 | 0.38965863 | 0.46293176 | 0 |
| HarmonicCentrality | 0.63197581 | 0.85001856 | 0.70172933 | 0.7279079 | 0 |
| Modularity | 5.07E-12 | 6.26E-14 | 2.05E-13 | 1.7778E-12 | 1 |
| PageRank | 0.95389978 | 0.00062006 | 0.49794577 | 0.4841552 | 0 |
| BP | 1.85E-05 | 8.25E-10 | 7.10E-08 | 6.1826E-06 | 1 |
| CC | 0.02245893 | 0.02912777 | 0.25275094 | 0.10144588 | 0 |
| MF | 0.00478777 | 8.65E-06 | 0.00401899 | 0.00293847 | 1 |
| FrequencyA | 0.97452688 | 0.65477107 | 0.75027966 | 0.79319254 | 0 |
| FrequencyC | 0.02572933 | 0.00511117 | 0.00818521 | 0.01300857 | 1 |
| FrequencyD | 0.04130471 | 0.09402434 | 0.04848896 | 0.06127267 | 0 |
| FrequencyE | 7.07E-08 | 4.34E-07 | 2.58E-10 | 1.6823E-07 | 1 |
| FrequencyF | 0.32605403 | 0.50984185 | 0.15225706 | 0.32938431 | 0 |
| FrequencyG | 0.00046053 | 0.0005 | 5.27E-05 | 0.00033773 | 1 |
| FrequencyH | 0.00389164 | 0.00798003 | 0.002501 | 0.00479089 | 1 |
| FrequencyI | 0.23199259 | 0.83295242 | 0.74675442 | 0.60389981 | 0 |
| FrequencyK | 6.72E-05 | 0.00050241 | 6.07E-10 | 0.00018989 | 1 |
| FrequencyL | 0.99361525 | 0.49701406 | 0.82531158 | 0.7719803 | 0 |
| FrequencyM | 0.61930436 | 0.54823941 | 0.77369498 | 0.64707958 | 0 |
| FrequencyN | 0.12300615 | 0.57499948 | 0.38479047 | 0.36093203 | 0 |
| FrequencyP | 0.04736054 | 0.02885271 | 0.03470955 | 0.03697426 | 1 |
| FrequencyQ | 0.08500627 | 0.1563478 | 0.09675047 | 0.11270151 | 0 |
| FrequencyR | 2.77E-05 | 0.00050473 | 2.35E-05 | 0.00018529 | 1 |
| FrequencyS | 0.00657788 | 0.25862913 | 3.71E-05 | 0.0884147 | 0 |
| FrequencyT | 0.00700462 | 0.07160317 | 4.06E-05 | 0.02621614 | 1 |
| FrequencyV | 0.62136512 | 0.73247842 | 0.91805705 | 0.7573002 | 0 |
| FrequencyW | 0.04223229 | 0.05276997 | 0.003008 | 0.03267009 | 1 |
| FrequencyX | 0.00170015 | 0.509777 | 0.28411294 | 0.2651967 | 0 |
| Aromaticity | 0.00555616 | 0.18419896 | 0.02105466 | 0.07026993 | 0 |
| Isoelectric | 0.66116371 | 0.86609529 | 0.33436489 | 0.62054129 | 0 |
| SSfractionHelix | 0.46595856 | 0.43765846 | 0.37168816 | 0.42510172 | 0 |
| SSfractionTurn | 4.03E-06 | 0.00026892 | 5.52E-08 | 9.1004E-05 | 1 |
| SSfractionSheet | 4.03E-06 | 0.00026892 | 5.52E-08 | 9.1004E-05 | 1 |

**Mann Whitney U Test**

Threshold is 0.05, if p-value < 0.05, then it is normal.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Dataset 1 | Dataset 2 | Dataset 3 | Minimum P-Value |
|
| BP | 1.56E-05 | 2.17E-12 | 5.12E-08 | 2.17E-12 |
| FrequencyE | 2.61E-08 | 1.71E-07 | 6.33E-11 | 6.33E-11 |
| Modularity | 6.49E-09 | 1.06E-10 | 1.48E-09 | 1.06E-10 |
| SSfractionTurn | 2.07E-05 | 6.90E-05 | 1.93E-08 | 1.93E-08 |
| SSfractionSheet | 2.07E-05 | 6.90E-05 | 1.93E-08 | 1.93E-08 |
| FrequencyK | 5.47E-04 | 0.00357163 | 4.03E-08 | 4.03E-08 |
| FrequencyG | 0.00119817 | 0.00081677 | 3.50E-06 | 0.0000035 |
| FrequencyS | 0.00280795 | 0.08700902 | 3.88E-06 | 0.00000388 |
| FrequencyT | 0.0005346 | 0.00997511 | 1.89E-05 | 0.0000189 |
| MF | 0.00095518 | 2.55E-05 | 0.00190748 | 0.0000255 |
| FrequencyW | 0.00226226 | 0.00620163 | 5.18E-05 | 0.0000518 |
| FrequencyR | 1.41E-04 | 0.00485175 | 7.37E-05 | 0.0000737 |
| FrequencyX | 0.00021659 | 0.11580676 | 0.045543 | 0.00021659 |
| FrequencyC | 0.02749799 | 0.00058427 | 0.00117648 | 0.00058427 |
| CC | 0.00179031 | 0.00164089 | 0.0332173 | 0.00164089 |
| FrequencyP | 0.04277138 | 0.01416375 | 0.00180336 | 0.00180336 |
| Aromaticity | 0.00305932 | 0.07495655 | 0.02184785 | 0.00305932 |
| FrequencyH | 0.01068879 | 0.00430951 | 0.01077978 | 0.00430951 |
| FrequencyD | 0.01299181 | 0.04616211 | 0.0063004 | 0.0063004 |
| PageRank | 0.4466116 | 0.01179905 | 0.12431628 | 0.01179905 |
| EigenvectorCentrality | 0.3432205 | 0.03354617 | 0.4357012 | 0.03354617 |
| FrequencyN | 0.04819396 | 0.27330352 | 0.14086425 | 0.04819396 |
| FrequencyQ | 0.0494191 | 0.13979546 | 0.1429538 | 0.0494191 |
| BetweennessCentrality | 0.06441373 | 0.063291 | 0.15322658 | 0.063291 |
| Isoelectric | 0.33520764 | 0.11217136 | 0.38185339 | 0.11217136 |
| ClosenessCentrality | 0.43497132 | 0.18173224 | 0.11419658 | 0.11419658 |
| FrequencyL | 0.3316665 | 0.45075692 | 0.14810413 | 0.14810413 |
| HarmonicCentrality | 0.16891062 | 0.20964212 | 0.31522333 | 0.16891062 |
| Average Shortest Path to all Disease genes | 0.39315089 | 0.17279292 | 0.22577515 | 0.17279292 |
| FrequencyF | 0.18394629 | 0.3423583 | 0.20524233 | 0.18394629 |
| SSfractionHelix | 0.19036167 | 0.25456843 | 0.29914535 | 0.19036167 |
| DegreeCentrality | 0.33789964 | 0.33359921 | 0.20598561 | 0.20598561 |
| FrequencyI | 0.20663299 | 0.49303442 | 0.43722691 | 0.20663299 |
| FrequencyM | 0.31607599 | 0.39387434 | 0.48062146 | 0.31607599 |
| FrequencyV | 0.33830483 | 0.32865012 | 0.45501189 | 0.32865012 |
| FrequencyA | 0.35426003 | 0.39813535 | 0.36342308 | 0.35426003 |

**Previous result**

|  |  |
| --- | --- |
| **Feature** | **p-value** |
| betweennessCentrality | 0.0 |
| degreeCentrality | 0.0 |
| eigenvectorCentrality | 0.0 |
| pageRank | 0.0 |
| frequencyC | 5.0711617337839556e-272 |
| bp | 2.2551136554012031e-266 |
| frequencyP | 3.1809961982607242e-205 |
| modularity | 7.7864781074210468e-197 |
| frequencyG | 3.0935046358956275e-169 |
| frequencyA | 5.6546162404219853e-168 |
| frequencyR | 1.5852180244653675e-167 |
| frequencyQ | 6.6957218016767694e-161 |
| frequencyW | 1.5991249980870714e-158 |
| mf | 7.7018610203437641e-156 |
| isoelectric | 3.7564798742695547e-154 |
| ssfractionTurn | 1.9318871410386582e-143 |
| ssfractionSheet | 1.9318871410386582e-143 |
| frequencyE | 5.4394821456192792e-143 |
| frequencyK | 3.2295339759509463e-135 |
| frequencyS | 6.0544780905567766e-133 |
| frequencyM | 3.2746933783387814e-129 |
| frequencyH | 1.7654241315856116e-119 |
| cc | 1.7797019013339253e-116 |
| frequencyY | 1.5314214346519035e-105 |
| frequencyF | 1.0239487524761442e-102 |
| frequencyT | 3.0786315358753929e-97 |
| frequencyD | 4.9152092390936651e-96 |
| frequencyL | 3.203509005186527e-87 |
| frequencyV | 2.3207878305453458e-86 |
| avgSP | 3.2491338206333851e-82 |
| aromaticity | 1.1898493070107014e-77 |
| frequencyN | 2.3340655068775652e-70 |
| frequencyI | 8.0232673908148838e-69 |
| ssfractionHelix | 2.567340502716567e-42 |
| harmonicCentrality | 1.6702656930961926e-23 |
| closenessCentrality | 3.9056997044935054e-19 |