**Supplementary Table 1.**

1.1 Comparison of VAGA\_CDA with other models.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Method | RWR-KNN | RWRLR | GCNCDA | DAERF | SGANRDA | GATNNCDA | CIRCWALK | **VAGA-CDA** |
| AUC | 0.9297 | 0.96 | 0.9090 | 0.9486 | 0.9411 | 0.9613 | 0.9777 | **0.9887** |

* Methods abbreviations are:- RWR-KNN: Random walk with restart and K-nearest algorithm; RWRLR: random walk with restart and Logistic regression; GCNCDA: Graph Convolutional Network Circular RNA-Disease Association; DAERF: Deep Auto encoder and Random Forest; SGANRDA: Semi-supervised Generative Adversarial Networks; GATNNCDA: Graph Attention Network and Multi-Layer Neural Network; CIRCWALK: Circular RNA Deep walk Algorithm; VAGA-CDA: Variational Autoencoder and Genetic algorithm for Circular RNA-Disease Association prediction.

1.2 Literature review on Circular RNA-Disease Association

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sl no. | Paper | Methodology | AUC | General Dataset | Observations |
| 1. | Integrating random walk with restart and k-Nearest Neighbour to identify novel circRNA-disease association[Xiujuan Lei et al, 2020]. | Random walk with restart  K-Nearest Neighbour | 0.9297 | CircR2Disease | * Disease semantic similarity, circRNA function similarity and GIP kernel similarity * Simple to implement * It is not applicable to new circRNAs or diseases without any known association |
| 2. | Predicting novel CircRNA-disease associations based on random walk and logistic regression model[Yulian Ding et al,2020]. | Random walk and logistic regression | 0.96 | CircR2Disease | * CircRNA function similarity * Simple to implement and fast running * It does not include enough information of diseases. It is not applicable to new circRNAs or diseases without any known association. |
| 3. | GCNCDA: A new method for predicting circRNA-disease associations based on Graph Convolutional Network Algorithm[Lei Wang et al, 2020]. | Graph Convolutional Network | 0.9090 | CircR2Disease | * GIP kernel similarity and disease semantic similarity * It effectively integrates network characteristics * GIP kernel similarity relies heavily on known circRNA-disease associations, thus causing inevitable bias. |
| 4. | Inferring Potential CircRNA–Disease Associations via Deep Autoencoder‑Based Classification , October [Deepti K et al ,2020]. | Deep Auto encoder,  Random forest | 0.9486 | CircR2Disease | * Integrated circRNA, disease similarities, and experimentally verifed circRNA–disease associations and extracted the hidden characteristics from them with the deep autoencoder. * Powerful computational tool for identifying novel circRNA–disease associations * The real negative samples were hard or even impossible to acquire. * Experimentally confrmed circRNA–disease relationships used in this study are far from enough. |
| 5. | SGANRDA: semi-supervised generative adversarial networks for predicting circRNA–disease associations[Lei Wang et al ,2021]. | semi-supervised generative adversarial networks | 0.9411 | CircR2Disease | * Integrates a variety of biological information. * Can identify new disease-associated circRNAs effectively. * Features extracted by natural language processing may not be obvious * Not easy to interpret the initialization and optimization process of hyperparameters in GAN network. * the randomly selected samples will bring some errors to the prediction results. |
| 6. | GATNNCDA: A Method Based on Graph Attention Network and Multi-Layer Neural Network for Predicting circRNA-Disease Associations[Cunmei Ji et al,2021]. | Graph Attention Network.  Multi-Layer Neural Network | 0.9613 | CircR2Disease | * Integrates circRNA functional similarity, disease semantic similarity and the GIP similarities. * linear transformation to project the integrated similarity matrices into the same space. * Applies a graph attention network to extract dense representations of nodes in the heterogeneous circRNA–disease graph. * The initial node features may not be perfect. * Known interactions between circRNA–disease associations are insufficient. |
| 7. | CircWalk: a novel approach to predict CircRNA‑disease association based on heterogeneous network representation learning [Morteza Kouhsar et al,2022]. | Deep Walk algorithm | 0.9777 | CircR2Disease  Mir2Disease | * Integrate multiple resources to construct a heterogeneous network from circRNAs, mRNAs, miRNAs, and diseases. * Each dataset uses unique identifiers and converting these identifiers sometimes can be impossible. * The lack of validated negative classes |