RAG Retrieval – Top-4 Chunks per Question (Annotated)

Q1: What is the main goal of feature selection in machine learning?

- Rank 1 | Score 0.7161 | Source data\Features_selection_1.pdf | Page 2 | Relevant: Yes

  computational time signi ficantly because the irrelevant and redundant features clutter the learning algorithm (Yu and Liu, 2004). Feature selection is a common way to minimize the problem of excessive and irrelevant features (Figure 2). Generally, feature selection methods reduce the dimensionality of the training data by excluding SNPs that: 1) have low or negligible predictive power for the phen...

- Rank 2 | Score 0.7525 | Source data\Features_selection_2.pdf | Page 11 | Relevant: No

  9. Conclusion In this paper we have tried to provide an introduction to feature selection techniques. The literature on feature selection techniques is very vast encompassing the applications of machine learning and pattern recognition. Comparison between feature selection algorithms can only be done using a single dataset since each underlying algorithm will behave differently for different data....

- Rank 3 | Score 0.7563 | Source data\Features_selection_1.pdf | Page 6 | Relevant: No

  while features that score above it are selected. Once a subset of features is selected, it can then be presented as an input to the chosen classifier algorithm. Unlike the other feature selection methods (wrapper and embedded), filter methods are independent/separate from the classi fier algorithm (Figure 5A). This separation means that filter methods are free from classifier’s bias which reduces overfi...

- Rank 4 | Score 0.7599 | Source data\Features_selection_1.pdf | Page 2 | Relevant: Yes

  feature selection) are typically assumed to be associated with loci that are mechanistically or functionally related to the underlying disease etiology (Pal and Foody, 2010; López et al., 2018). Therefore, extracting a subset of the most relevant features (through feature selection) could help researchers to understand the biological process(es) that underlie the disease (Cueto-López et al., 2019)...

Q2: What are the three classical categories of feature selection methods?

- Rank 1 | Score 0.6240 | Source data\Features_selection_1.pdf | Page 9 | Relevant: No

  New feature selection strategies are emerging that either: 1), use a two-step strategy with a combination of different feature selection methods (hybrid methods); or 2), combine the output of multiple feature selection methods (ensemble methods). These strategies take advantage of the strengths of the different feature selection methods that they include. 3 HYBRID METHODS—COMBINING DIFFERENT FEATU...

- Rank 2 | Score 0.6267 | Source data\Features_selection_1.pdf | Page 5 | Relevant: No

  the potential to detect redundancies and interactions between features. The particular strengths and weaknesses of each methodological category mean they are more suitable for particular use cases (Saeys et al., 2007; Okser et al., 2013; De et al., 2014; Remeseiro and Bolon-Canedo, 2019)(Table 1). 2.1 Filter Methods for Feature Selection Filter methods use feature ranking as the evaluation metric...

- Rank 3 | Score 0.6999 | Source data\Features_selection_1.pdf | Page 4 | Relevant: Yes

  2 FEATURE SELECTION TECHNIQUES The feature selection methods that are routinely used in classification can be split into three methodological categories (Guyon et al., 2008; Bolón-Canedo et al., 2013): 1)filters; 2) wrappers; and 3) embedded methods (Table 1). These methods differ in terms of 1) the feature selection aspect being separate or integrated as a part of the learning algorithm; 2) evaluat...

- Rank 4 | Score 0.7135 | Source data\Features_selection_1.pdf | Page 8 | Relevant: No

  Instead, interaction terms must be explicitly included in the analysis (Signorino and Kirchner, 2018). This is commonly achieved by exhaustively including all (usually pairwise) interaction terms for the features. While this approach can be effective for data with low dimensionality, it can be inaccurate and computationally prohibitive in highly dimensional data settings. Two-stage or hybrid strat...

Q3: How do filter methods select features, and what is their main advantage?

- Rank 1 | Score 0.5213 | Source data\Features_selection_1.pdf | Page 6 | Relevant: Yes

  while features that score above it are selected. Once a subset of features is selected, it can then be presented as an input to the chosen classifier algorithm. Unlike the other feature selection methods (wrapper and embedded), filter methods are independent/separate from the classifier algorithm (Figure 5A). This separation means that filter methods are free from classifier's bias which reduces overfi...

- Rank 2 | Score 0.6386 | Source data\Features_selection_1.pdf | Page 5 | Relevant: Yes

  the potential to detect redundancies and interactions between features. The particular strengths and weaknesses of each methodological category mean they are more suitable for particular use cases (Saeys et al., 2007; Okser et al., 2013; De et al., 2014; Remeseiro and Bolon-Canedo, 2019)(Table 1). 2.1 Filter Methods for Feature Selection Filter methods use feature ranking as the evaluation metric...

- Rank 3 | Score 0.6453 | Source data\Features_selection_1.pdf | Page 9 | Relevant: No

  New feature selection strategies are emerging that either: 1), use a two-step strategy with a combination of different feature selection methods (hybrid methods); or 2), combine the output of multiple feature selection methods (ensemble methods). These strategies take advantage of the strengths of the different feature selection methods that they include. 3 HYBRID METHODS—COMBINING DIFFERENT FEATU...

- Rank 4 | Score 0.7587 | Source data\Features_selection_1.pdf | Page 9 | Relevant: No

  for all problem settings (Wolpert and Macready, 1997). This is true for feature selection methods, each of which has its own strengths and weaknesses (Table 1), relying on different metrics and underlying assumptions. Several studies have compared the predictive performance of the different feature selection methods (Forman, 2003; Bolón-Canedo et al., 2013; Aphinyanaphongs et al., 2014; Wah et al....

Q4: What characterizes wrapper methods, and what is their main drawback?

- Rank 1 | Score 0.8265 | Source data\Features_selection_1.pdf | Page 7 | Relevant: Yes

  training dataset, but poor generalizability to external datasets) (i.e., more prone to overfitting) (Kohavi and John, 1997). Unlike filter methods which produce a ranked list of features, wrapper methods produce a"best" feature subset as the output. This has both advantages and disadvantages. One advantage of this is that the user does not need to determine the most optimum threshold or number of fe...

- Rank 2 | Score 1.0060 | Source data\Features_selection_1.pdf | Page 8 | Relevant: No

  methods mentioned above also return a ranked list of features. Decision tree-based algorithms rank feature importance based on metrics like the Mean Decrease Impurity (MDI) (Louppe et al., 2013). For regularization methods, the ranking of features is provided by the magnitude of the feature coefficients. Embedded methods are an intermediate solution between filter and wrapper methods in the sense th...

- Rank 3 | Score 1.0425 | Source data\Features_selection_2.pdf | Page 4 | Relevant: Yes

  tions with weighting/penalty imposing characteristics. A binary PSO[32,48,4] algorithm can also be used for wrapper imple- mentation. In[49] comparison between GA and PSO using SVM for gene selection can be found. The main drawback of Wrapper methods is the number of computations required to obtain the feature subset. For each subset evaluation, the predictor creates a new model i.e. the predictor...

- Rank 4 | Score 1.0939 | Source data\Features_selection_2.pdf | Page 2 | Relevant: Yes

  3. Wrapper methods Wrapper methods use the predictor as a black box and the predictor performance as the objective function to evaluate the variable subset. Since evaluating 2N subsets becomes a NP-hard problem, suboptimal subsets are found by employing search algorithms which find a subset heuristically. A number of search algorithms can be used to find a subset of variables which maximizes the obj...

Q5: What is the key idea behind embedded methods such as mRMR (max-relevancy, min-redundancy)?

- Rank 1 | Score 0.9388 | Source data\Features_selection_2.pdf | Page 5 | Relevant: Yes

  method. The mRMR (max-relevancy, min-redundancy)[24] is another method based on MI. It uses similar criteria as in(7) given as: Iðxj; CÞ/C0 1 m /C0 1 X xl2Sm/C0 1 Iðxj; xlÞ ð8Þ where xi is themth feature in subsetS and the setSm/C0 1 is the so far selected subset withm /C0 1 features. Instead of a greedy algorithm a two stage approach is implemented. First the criterion(8) is used to select a numb...

- Rank 2 | Score 1.0460 | Source data\Features_selection_2.pdf | Page 5 | Relevant: Yes

  objective function is designed such that choosing a feature will maximize the MI between the feature and the class output while the MI between the selected feature and the subset of the so far selected features is a minimum. This is formulated as: IðY; f Þ/C0 b X s2S Iðf ; sÞ ð7Þ where Y is the output,f is the current selected feature,s is the feature in the already selected subsetS and b controls...

- Rank 3 | Score 1.0464 | Source data\Features_selection_1.pdf | Page 8 | Relevant: No

  methods mentioned above also return a ranked list of features. Decision tree-based algorithms rank feature importance based on metrics like the Mean Decrease Impurity (MDI) (Louppe et al., 2013). For regularization methods, the ranking of features is provided by the magnitude of the feature coefficients. Embedded methods are an intermediate solution between filter and wrapper methods in the sense th...

- Rank 4 | Score 1.1509 | Source data\Features_selection_1.pdf | Page 7 | Relevant: Yes

  Selection In an embedded method, feature selection is integrated or built into the classifier algorithm. During the training step, the classifier adjusts its internal parameters and determines the appropriate weights/importance given for each feature to produce the best classification accuracy. Therefore, the search for the optimum feature subset and model construction in an embedded method is combi...

Q6: What is the Pearson correlation coefficient used for in feature selection?

- Rank 1 | Score 0.9244 | Source data\Features_selection_2.pdf | Page 2 | Relevant: No

  not discriminate the variables in terms of the correlation to other variables. The variables in the subset can be highly corre- lated in that a smaller subset would suffice[11,28]. This issue of redundant vs. relevant variables is addressed in[1] with good examples. In feature ranking, important features that are less informative on their own but are informative when com- bined with others could be...

- Rank 2 | Score 0.9770 | Source data\Features_selection_3.pdf | Page 4 | Relevant: Yes

  for the average dispersion within class and the average scatter dis- tance among classes. The value of the judgment function in each feature subset is calculated, and the feature subset that maximizes the judgment function value is selected to determine the candidate feature. Finally, the correlation coeficient between the candidate feature and the selected feature is calculated. If the correl...

- Rank 3 | Score 0.9844 | Source data\Features_selection_2.pdf | Page 0 | Relevant: No

  ing tasks. Feature Selection (variable elimination) helps in understanding data, reducing computation requirement, reducing the effect of curse of dimensionality and improving the predictor performance. In this paper we look at some of the methods found in literature which use particular measurements to find a subset of variables (features) which improves the overall prediction performance. The foc...

- Rank 4 | Score 1.0585 | Source data\Features_selection_2.pdf | Page 0 | Relevant: No

  which many of them could be highly correlated with other variables (e.g. when two features are perfectly correlated, only one feature is sufficient to describe the data). The dependant variables provide no extra information about the classes and thus serve as noise for the predictor. This means that the total information content can be obtained from fewer unique fea- tures which contain maximum dis...

Q7: How is Mutual Information (MI) interpreted in the context of feature selection?

- Rank 1 | Score 0.6725 | Source data\Features_selection_3.pdf | Page 7 | Relevant: No

  [82] J. Novovicová, P. Somol , M. Haindl , P. Pudil , Conditional mutual informa- tion based feature selection for classification task, in: Proceedings of the 12th Iberoamerican Conference on Congress on Pattern Recognition, 2007, pp. 417–426 . [83] Y. Zhang , Z. Zhang , Feature subset selection with cumulate conditional mutual information minimization, Expert Syst. Appl. 39 (2012) 6078–6088...

- Rank 2 | Score 0.6856 | Source data\Features_selection_3.pdf | Page 2 | Relevant: Yes

  ever, mutual information based MRMR only minimizes feature– feature mutual information and ignores the classification perfor- mance of candidate features, which might be influenced by the selected features. Conditional mutual information analysis is then introduced to overcome this problem. Feature selection methods based on conditional mutual information have attracted significant attention [43...

- Rank 3 | Score 0.7497 | Source data\Features_selection_3.pdf | Page 2 | Relevant: Yes

  expression of relevance analysis, and the later item belongs to re- dundancy analysis. This expression can be further extended into conditional mutual information form shown in formula (11) . MRMR : max [1 | S| ∑ x i ∈ S I(x i ;C) − 1 | S | 2 ∑ x i ∈ S ∑ x j ∈ S I (x i ;x j)] (10) CMRMR : (11) max [1 | S| ∑ x i ∈ S I(x i ;C) − 1 | S | 2 ∑ x i ∈ S ∑ x j ∈ S I (x i ;xj)] ...

- Rank 4 | Score 0.7853 | Source data\Features_selection_1.pdf | Page 6 | Relevant: No

  each feature is considered separately, univariate methods only focus on feature relevance and cannot detect feature redundancy, or interactions. This decreases model predictor performance because: 1) the inclusion of redundant features makes the feature subset larger than necessary; and 2) ignoring feature interactions can lead to the loss of important information. More advanced multivariate filter...

Q8: What is Sequential Forward Selection (SFS) and how does it work?

- Rank 1 | Score 0.5765 | Source data\Features_selection_2.pdf | Page 3 | Relevant: Yes

  racy. The process is repeated until the required number of features are added. This is a naive SFS algorithm since the depen- dency between the features is not accounted for. A Sequential Backward Selection (SBS) algorithm can also be constructed which is similar to SFS but the algorithm starts from the complete set of variables and removes one feature at a time whose removal gives the lowest decr...

- Rank 2 | Score 0.8493 | Source data\Features_selection_2.pdf | Page 4 | Relevant: No

  The SFS and SFFS methods suffer from producing nested subsets since the forward inclusion was always unconditional which means that two highly correlated variables might be included if it gave the highest performance in the SFS evaluation. To avoid the nesting effect, adaptive version of the SFFS was developed in[35,36]. The Adaptive Sequential Forward Floating Selection (ASFFS) algorithm used a p...

- Rank 3 | Score 0.9688 | Source data\Features_selection_2.pdf | Page 3 | Relevant: No

  ferent subsets to optimize the objective function. Different subsets are generated either by searching around in a search- space or by generating solutions to the optimization problem. First we will look at sequential selection algorithms followed by the heuristic search algorithms. 3.1. Sequential selection algorithms These algorithms are called sequential due to the iterative nature of the algor...

- Rank 4 | Score 1.1240 | Source data\Features_selection_2.pdf | Page 3 | Relevant: No

  on the objective function. The SFFS algorithm adds another step which excludes one feature at a time from the subset ob- tained in the first step and evaluates the new subsets. If excluding a feature increases the value of the objective function then that feature is removed and goes back to the first step with the new reduced subset or else the algorithm is repeated from the top. This process is rep...

Q9: What problem does Sequential Floating Forward Selection (SFFS) try to solve compared to naive SI

- Rank 1 | Score 0.4431 | Source data\Features_selection_2.pdf | Page 3 | Relevant: No

  racy. The process is repeated until the required number of features are added. This is a naive SFS algorithm since the depen- dency between the features is not accounted for. A Sequential Backward Selection (SBS) algorithm can also be constructed which is similar to SFS but the algorithm starts from the complete set of variables and removes one feature at a time whose removal gives the lowest decr...

- Rank 2 | Score 0.4972 | Source data\Features_selection_2.pdf | Page 4 | Relevant: No

  The SFS and SFFS methods suffer from producing nested subsets since the forward inclusion was always unconditional which means that two highly correlated variables might be included if it gave the highest performance in the SFS evaluation. To avoid the nesting effect, adaptive version of the SFFS was developed in[35,36]. The Adaptive Sequential Forward Floating Selection (ASFFS) algorithm used a p...

- Rank 3 | Score 0.9424 | Source data\Features_selection_2.pdf | Page 3 | Relevant: Yes

  on the objective function. The SFFS algorithm adds another step which excludes one feature at a time from the subset ob- tained in the first step and evaluates the new subsets. If excluding a feature increases the value of the objective function then that feature is removed and goes back to the first step with the new reduced subset or else the algorithm is repeated from the top. This process is rep...

- Rank 4 | Score 0.9976 | Source data\Features_selection_2.pdf | Page 4 | Relevant: No

  in [9,10,33,35]. Theoretically, the ASFFS should produce a better subset than SFFS but this is dependent on the objective func- tion and the distribution of the data. The Plus-L-Minus-r search method[35,37,38] also tries to avoid nesting. In the Plus-L- Minus-r search, in each cycleL variables were added andr variables were removed until the desired subset was achieved. The parametersL and r have ...

Q10: According to Cai et al. (2018), how are supervised, unsupervised, and semi-supervised feature sel

- Rank 1 | Score 0.5878 | Source data\Features_selection_3.pdf | Page 5 | Relevant: Yes

  J. Cai et al. / Neurocomputing 300 (2018) 70–79 75 5. Semi-supervised feature selection Given the dataset D = { D l , D u }, where D l is the sample set with class labels, and D u is the sample set without class labels, semi- supervised learning model uses D u to improve the learning per- formance of learning model trained by D l . Semi-supervised feature selection methods, which are mainly ...

- Rank 2 | Score 0.5919 | Source data\Features_selection_2.pdf | Page 6 | Relevant: Yes

  5. Other feature selection techniques Unsupervised learning deals with finding hidden structure in unlabelled data. Clustering techniques[7] are a primary example of unsupervised learning which tries to discover natural groupings in a set of objects without knowledge of class labels. Feature selection using unsupervised learning techniques are beyond the scope of this paper and will not be discusse...

- Rank 3 | Score 0.6094 | Source data\Features_selection_3.pdf | Page 5 | Relevant: No

  The methods mentioned above are usually scored against in- dividual features. Existing researches emphasize the redundancy analysis among features in designing the semi-supervised feature selection methods [140–142] . Benabdeslem et al proposed a filter approach based on a constrained Laplacian score, in which the re- dundancy is removed after the relevant features are selected [140] . Wang p...

- Rank 4 | Score 0.6115 | Source data\Features_selection_2.pdf | Page 6 | Relevant: Yes

  Semi-supervised learning is another class wherein both labelled and unlabelled data are used for learning[7,29,66,67].I t uses both labelled data (less number of samples) and unlabelled data (abundantly available) to modify a hypothesis obtained from labelled data alone. In[67] the authors use a clustering indicator construction to score a set of features. In[29] the authors use the maximum margin...

Q11: What are “hybrid” and “ensemble” feature selection strategies mentioned in the survey?

- Rank 1 | Score 0.5838 | Source data\Features_selection_1.pdf | Page 11 | Relevant: Yes

  shown that ensemble feature selection methods tend to produce better classification accuracy than is achieved using single feature selection methods (Seijo-Pardo et al., 2015; Hoque et al., 2017; Wang et al., 2019; Tsai and Sung, 2020). Furthermore, ensemble feature selection can improve the stability of the selected feature set (i.e., it is more robust to small changes in the input data) (Yang and...

- Rank 2 | Score 0.6122 | Source data\Features_selection_1.pdf | Page 8 | Relevant: Yes

  FIGURE 6 | (A)Generalized illustration of ensemble methods. In ensemble methods, the outputs of several feature selection methods are aggregated to obtain the final selected features. FS = feature selection.(B) Generalized illustration of majority voting system where the different generated feature subsets are used to train and test a specific classifier. The final output is the class predicted by the...

- Rank 3 | Score 0.6217 | Source data\Features_selection_1.pdf | Page 8 | Relevant: Yes

  Instead, interaction terms must be explicitly included in the analysis (Signorino and Kirchner, 2018). This is commonly achieved by exhaustively including all (usually pairwise) interaction terms for the features. While this approach can be effective for data with low dimensionality, it can be inaccurate and computationally prohibitive in highly dimensional data settings. Two-stage or hybrid strat...

- Rank 4 | Score 0.6532 | Source data\Features_selection_1.pdf | Page 11 | Relevant: Yes

  biological knowledge are incomplete. Therefore, relying on external a priori knowledge will hinder the identification of novel variants outside our current biological understanding. 3.2 Ensemble Method—Combining the Output of Different Feature Selections Ensemble feature selection methods are based on the assumption that combining the output of multiple algorithms is better than using the output of...

Q12: Why is the stability of feature selection algorithms an important issue?

- Rank 1 | Score 0.6997 | Source data\Features_selection_1.pdf | Page 11 | Relevant: Yes

  shown that ensemble feature selection methods tend to produce better classification accuracy than is achieved using single feature selection methods (Seijo-Pardo et al., 2015; Hoque et al., 2017; Wang et al., 2019; Tsai and Sung, 2020). Furthermore, ensemble feature selection can improve the stability of the selected feature set (i.e., it is more robust to small changes in the input data) (Yang and...

- Rank 2 | Score 0.7181 | Source data\Features_selection_2.pdf | Page 6 | Relevant: Yes

  aggregation methods such as ensemble-mean, linear aggregation, weighted aggregation methods to obtain the final feature subset. 6. Stability of feature selection algorithms For a particular application, various feature selection algorithms can be applied and the best one can be selected which meets the required criteria. An overlooked problem is the stability of the feature selection algorithms. St...

- Rank 3 | Score 0.7250 | Source data\Features_selection_2.pdf | Page 12 | Relevant: No

  [68] Haury A-C, Gestraud P, Vert J-P. The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. PLoS ONE 2011;6:e28210. [69] T A, T H, de Peer Y V, P D, Y S. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. Bioinformatics 2010;26:392–8. [70] Dunne K, Cunningham P, Azuaje F. Solutions to instability pro...

- Rank 4 | Score 0.7459 | Source data\Features_selection_1.pdf | Page 2 | Relevant: No

  computational time significantly because the irrelevant and redundant features clutter the learning algorithm (Yu and Liu, 2004). Feature selection is a common way to minimize the problem of excessive and irrelevant features (Figure 2). Generally, feature selection methods reduce the dimensionality of the training data by excluding SNPs that: 1) have low or negligible predictive power for the phen...

Q13: In Khan et al. (2020), what is the practical effect of using feature importance for the UNSW-NB15 ir

- Rank 1 | Score 0.6975 | Source data\Features_selection_4.pdf | Page 3 | Relevant: No

  The UNSW-NB 15 dataset includes complex patterns compared to NSL KDD99 dataset and contains 9 different attack types unlike NSL KDD99 dataset with only 5 different attack types. Hence for the analysis, UNSW-NB 15 dataset is used to evaluate various classification methods. 3.2 Data Preprocessing Feature Scaling. Standardization of a dataset is very important for many machine learning algorithms whi...

- Rank 2 | Score 0.7094 | Source data\Features_selection_4.pdf | Page 8 | Relevant: No

  Miskolc, Hungary (2018) 2. Kumar, K., Batth, J.S.: Network intrusion detection with feature selection techniques using machine-learning algorithms. Int. J. Comput. Appl. 150(12), 1 –13 (2016) 3. Belavagi, M.C., Muniyal, B.: Performance evaluation of supervised machine learning algorithms for intrusion detection. Procedia Comput. Sci. 89, 117 –123 (2016) 4. Mogal, D.G., Ghungrad, S.R., Bapusaheb, B...

- Rank 3 | Score 0.7659 | Source data\Features_selection_4.pdf | Page 6 | Relevant: Yes

  In (Fig. 2), X-axis represents the attribute names and Y-axis represents the level of importance of attributes. The 41 attributes were reduced to 11 attributes based on their importance by using feature importance on Random Forest classifier. The reduced features of UNSW-NB15 dataset are Service, State, Sbytes, Dbytes, Rate, Sttl, Ackdat, ct_dst_ltm, ct_src_dport_ltm, ct_dst_sport_ltm, and ct_src_...

- Rank 4 | Score 0.8500 | Source data\Features_selection_4.pdf | Page 0 | Relevant: Yes

  nique. These classifiers are chosen as they perform superior to other base and ensemble machine learning techniques after feature selection. Feature Impor- tance technique is utilized to obtain the highest ranked features. Reduced attributes improve the accuracy as well as decrease the computation time and prediction time. The experimental results on UNSW-NB dataset show that there is a drastic de...

Q14: Which classifiers are commonly used in the articles to evaluate the performance of feature selecti

- Rank 1 | Score 0.6832 | Source data\Features_selection_1.pdf | Page 5 | Relevant: No

  the potential to detect redundancies and interactions between features. The particular strengths and weaknesses of each methodological category mean they are more suitable for particular use cases (Saeys et al., 2007; Okser et al., 2013; De et al., 2014; Remeseiro and Bolon-Canedo, 2019)(Table 1). 2.1 Filter Methods for Feature Selection Filter methods use feature ranking as the evaluation metric...

- Rank 2 | Score 0.7009 | Source data\Features_selection_1.pdf | Page 11 | Relevant: No

  shown that ensemble feature selection methods tend to produce better classification accuracy than is achieved using single feature selection methods (Seijo-Pardo et al., 2015; Hoque et al., 2017; Wang et al., 2019; Tsai and Sung, 2020). Furthermore, ensemble feature selection can improve the stability of the selected feature set (i.e., it is more robust to small changes in the input data) (Yang and...

- Rank 3 | Score 0.7042 | Source data\Features_selection_2.pdf | Page 11 | Relevant: No

  9. Conclusion In this paper we have tried to provide an introduction to feature selection techniques. The literature on feature selection techniques is very vast encompassing the applications of machine learning and pattern recognition. Comparison between feature selection algorithms can only be done using a single dataset since each underlying algorithm will behave differently for different data....

- Rank 4 | Score 0.7305 | Source data\Features_selection_1.pdf | Page 9 | Relevant: No

  New feature selection strategies are emerging that either: 1), use a two-step strategy with a combination of different feature selection methods (hybrid methods); or 2), combine the output of multiple feature selection methods (ensemble methods). These strategies take advantage of the strengths of the different feature selection methods that they include. 3 HYBRID METHODS—COMBINING DIFFERENT FEATU...

Q15: According to the experiments in the survey paper, why can filter methods give “irregular” perform

- Rank 1 | Score 1.0428 | Source data\Features_selection_1.pdf | Page 6 | Relevant: No

  while features that score above it are selected. Once a subset of features is selected, it can then be presented as an input to the chosen classifier algorithm. Unlike the other feature selection methods (wrapper and embedded), filter methods are independent/separate from the classifier algorithm (Figure 5A). This separation means that filter methods are free from classifier’s bias which reduces overfi...

- Rank 2 | Score 1.1053 | Source data\Features_selection_1.pdf | Page 5 | Relevant: No

  the potential to detect redundancies and interactions between features. The particular strengths and weaknesses of each methodological category mean they are more suitable for particular use cases (Saeys et al., 2007; Okser et al., 2013; De et al., 2014; Remeseiro and Bolon-Canedo, 2019)(Table 1). 2.1 Filter Methods for Feature Selection Filter methods use feature ranking as the evaluation metric...

- Rank 3 | Score 1.1250 | Source data\Features_selection_1.pdf | Page 9 | Relevant: No

  New feature selection strategies are emerging that either: 1), use a two-step strategy with a combination of different feature selection methods (hybrid methods); or 2), combine the output of multiple feature selection methods (ensemble methods). These strategies take advantage of the strengths of the different feature selection methods that they include. 3 HYBRID METHODS—COMBINING DIFFERENT FEATU...

- Rank 4 | Score 1.1510 | Source data\Features_selection_1.pdf | Page 6 | Relevant: No

  or embedded methods. The main advantage of filter methods over other feature selection methods is that they are generally less computationally demanding, and thus can easily be scaled to very high dimensional data (e.g. SNP genotype datasets). Existing fillter methods can be broadly categorized as either univariate or multivariate. Univariate methods test each feature individually, while multivariate ...

Evaluation Summary

Total questions: 15
Total chunks (top-4 per question): 60
Total relevant chunks: 27
Recall@4 (questions with ≥1 relevant chunk): 0.867
Precision@4 (relevant chunks / total chunks): 0.450