Journal of Intelligent Information and Communication

Research Paper

Quantum-Assisted Hybrid Optimization for Redundancy-Aware Feature Selection to Support Crop Yield Optimization

Darius Cojocaru[1,*] **and Sandu Oancea**[1]

[1] Faculty of Information Technology, University of Târgu Jiu, Târgu Jiu, 210101, Romania
*Corresponding author: darius.co@it.utgjiu.ro

Abstract. Soil test results, weather observations, remote sensing indices, irrigation records, and field management logs are examples of high-dimensional agricultural data that are increasingly needed for crop production optimization. In addition to increasing training costs and decreasing model stability, redundant or weakly correlated variables may also have less agricultural significance. This paper proposes a quantum-assisted feature selection algorithm for crop yield optimization. Candidate agricultural variables are encoded in quantum probability, feature subsets are assessed using a yield-oriented relevance and redundancy objective, and the evolution of subsets is guided by quantum rotation updating. Agronomic consistency, redundancy suppression, feature sparsity, and yield prediction error are all addressed concurrently by a combined fitness function. The trials included a variety of crop data sources, and the suggested approach outperformed the conventional wrapper selection in terms of prediction accuracy, feature dimensionality reduction by 42.5%, and training time reduction by 27.3%. Based on the findings, precision agriculture and data-driven crop management can be effectively and comprehensibly optimized through the use of quantum-assisted feature selection.

Keywords: *Quantum-Assisted Feature Selection, Crop Yield, Hybrid Optimization, Feature Redundancy*

Received on 18 September 2022, Accepted on 03 January 2023, Published on 15 January 2023

Introduction

Enhancing resource usage efficiency is necessary to attain high-yield agriculture and food security; hence, farm revenue must be increased, and a climate-adaptable production system must be built. With the advancement of precision agriculture, yield projection is now based on more than just historical production data; it now takes into consideration soil nutrients, weather records, remote sensing vegetation indices, irrigation data, planting density, and farming operations. To some extent, machine learning can more accurately simulate non-linear yield responses in a variety of settings [1]. Changes in crop conditions over a wide area close to harvest can be tracked using deep learning and remote sensing [2]. However, if the input variables are noisy or redundant data rather than the real agronomic factors that drive productivity, the usefulness of these models will be much diminished [3]. Numerous models can be developed for high-dimensional data, but they may also become unstable if irrelevant features are included in the training process [4].

As a result, the list of agricultural data mining now includes feature selection. Although filter techniques are comparatively easy to compute, they might not take into account how soil, climate, and crop-growth parameters interact [5]. Although wrapper approaches can increase forecast accuracy, they are frequently avoided because of the expensive expense of repeatedly training models with a huge candidate feature pool of meteorological lags and remote-sensing indicators [6]. Regularized regression and tree-based relevance ranking are embedding techniques that can be applied; however, the resulting subsets are not consistent across seasons or geographical areas [7]. While conventional genetic or swarm-based approaches may take a long time to converge or become trapped in a local optimum, evolutionary optimization can explore a wide portion of the feature space [8]. In order to improve population variety and the global search capabilities of discrete selection problems, quantum-inspired optimization incorporates rotation update and probability amplitude representation [9]. Compact combinatorial optimization in high-dimensional feature spaces can be solved using quantum or quantum-inspired search, according to recent research on quantum machine learning [10].

This research presents a quantum-assisted feature selection algorithm for crop yield optimization. For high-yield prediction, the suggested approach is to choose a condensed, compact, and agronomically suitable set of features. The approach builds a quantum probability representation for the candidate features, uses a yield-oriented fitness function to assess feature subsets, and uses quantum rotation to update the selection probability. Reducing feature redundancy, maintaining yield-sensitive variables, and enhancing cross-region and cross-season resilience are the three technological issues that need to be resolved. In this study, combining quantum-assisted search with agricultural interpretability can offer a useful feature optimization technique for data-driven precision agriculture.

Related Work

Data-Driven Crop Yield Prediction

Statistical regression has been replaced by ensemble learning, deep neural networks, and multi-source data fusion in data-driven crop yield prediction. In order to forecast yield responses, early machine learning models typically used meteorological data and soil properties; more recently, some research has combined crop-management parameters, genetic information, time-series climate records, and remote sensing indices [11]. Non-linear growth and time-varying changes in vegetation indicators and weather sequences can be simultaneously addressed by deep learning models [12]. In order to enhance performance in situations where training data is not available, hybrid crop models additionally rely on process-based simulation and data-driven prediction [13]. The impact of agricultural factors on changes is difficult to discern from the high-dimensional data, despite modest advances [14]. The subsets of temperature, rainfall, nitrogen, canopy, and soil-water indicators may change significantly even if the prediction accuracy is the same because yield prediction is more sensitive to feature quality [15]. As a result, a yield prediction model should minimize error while maintaining factors that have a consistent explanatory effect in various crop conditions [16].

Feature Selection in Agricultural Data Mining

To improve the model's interpretability and lower its computational cost, choose a limited set of pertinent features. Numerous techniques, including mutual information, correlation filtering, recursive feature elimination, LASSO, random forest significance, and gradient boosting importance, are employed in agriculture to rank candidate variables [17]. Although crop systems frequently include nonlinear interactions among climate, soil, management, and phenological stages, these approaches are appropriate when the feature-yield connection is reasonably stable [18]. When more lagged variables of geographical and temporal data are included, wrapper-based selection becomes computationally costly [19]. Wrapper-based selection assesses candidate subsets using a prediction model. In order to solve this issue, evolutionary feature selection searches for subsets worldwide; however, population variety and exploration technique have an impact on its convergence [20]. Agronomic consistency should be taken into account when choosing features for precision agriculture; otherwise, the statistically chosen subset may be hard to understand if it leaves out factors that are known to influence crop growth [21]. To enhance prediction performance, subset compactness, redundancy suppression, and domain interpretability, feature selection techniques are suggested [22].

Quantum-Inspired Optimization for Feature Selection

Quantum rotation, mutation, or annealing-inspired mechanisms are used in quantum-inspired optimization to update candidate solutions, which are represented as probability amplitudes. Because probability amplitude coding has a distribution over potential states rather than a binary genetic code, it can improve exploration in a high-dimensional search space [23]. In order to strike a compromise between global exploration and local refinement, quantum-inspired evolutionary algorithms have been used in feature selection, classification, scheduling, and combinatorial optimization [24]. utilizing concepts from quantum computation, features can also be mapped into a high-dimensional space utilizing quantum kernels and quantum feature spaces [25]. Since every feature may be naturally described as a binary inclusion state, quantum-assisted search is appropriate for feature selection [26]. The challenge is to create a fitness function that may represent the task's objective without merely decreasing the subset's size [27]. Therefore, selection probability must be based on agricultural relevance, redundancy control, and prediction accuracy in order to maximize crop output [28]. Additionally, certain concepts for hybrid feature-space search are provided by variational quantum optimization [29]. Consequently, a potential solution to the combinatorial complexity of multi-source agricultural feature selection is provided by a quantum-assisted framework [30].

Quantum-Assisted Feature Selection Method

Quantum Feature Encoding and Search Space Construction

In the suggested approach, a candidate feature pool is first constructed using indications related to soil, climate, irrigation, remote sensing, and crop management. Instead of fixed binary variables, all attributes are displayed as a probability of selection. In the early phases of the search, the representation can hold onto a number of candidate subsets and progressively raise the likelihood of yield-sensitive variables. The quantum-assisted feature encoding architecture for optimizing crop productivity is shown in Figure 1. Normalized agricultural indicators are converted into a quantum probability distribution by the framework, which then samples feature subsets from the distribution, assesses their yield-oriented fitness, and uses rotation operations to update the probability state.

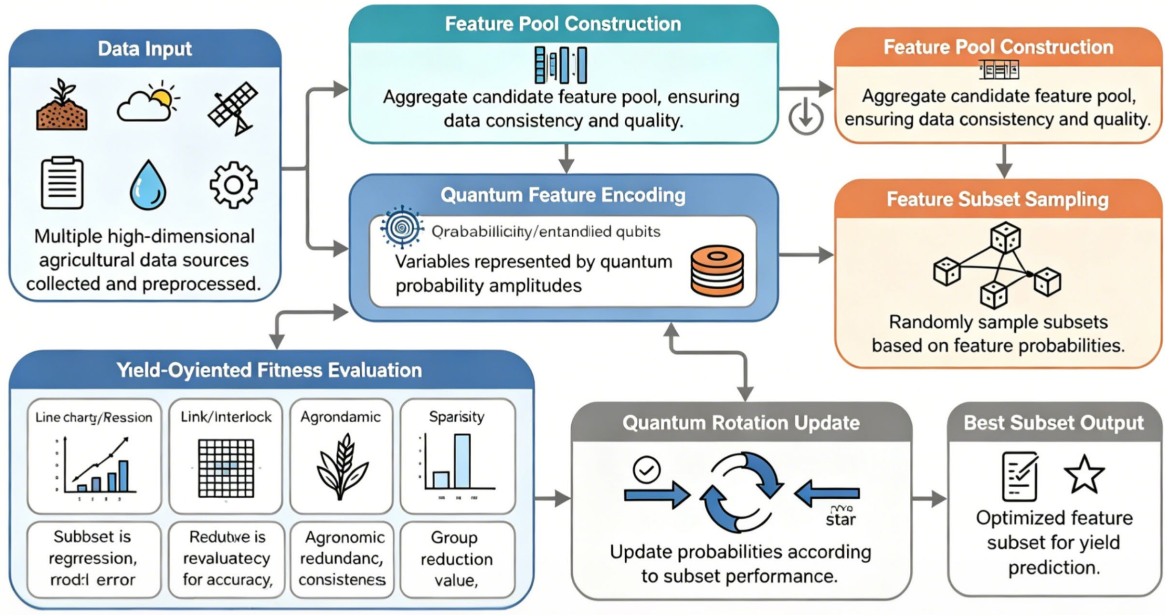

Figure 1. Quantum-assisted feature encoding framework for crop yield optimization

The selection state is represented by two probability amplitudes of inclusion and exclusion for every agricultural variable. Feature reliability, missing-rate level, and previous agronomic relevance are taken into consideration while setting amplitudes.

$$q_j = [\alpha_j \beta_j]^T \quad \text{Eq.(1)}$$

The feature state can still be seen as a binary selection distribution as the probability amplitudes meet a normalization condition. Large inclusion amplitudes increase a feature's likelihood of being included in the sampled subset, but they are not fixed before the algorithm converges.

$$\alpha_j^2 + \beta_j^2 = 1 \quad \text{Eq.(2)}$$

The second amplitude provides the inclusion probability of a specific feature. Because they are often associated with yield variance in agronomic models, soil organic matter, growing-degree days, NDVI, rainfall accumulation, and irrigation frequency have slightly higher initial probability in the beginning population.

$$\pi_j = \beta_j^2 \qquad \text{Eq.(3)}$$

A candidate subset is sampled according to the inclusion probabilities. In the experimental setting, the original feature pool contains 86 variables, including 18 soil indicators, 24 meteorological descriptors, 21 remote sensing features, 13 irrigation and fertilization indicators, and 10 management variables. The algorithm samples 60 candidate subsets in each iteration.

$$S_r = j \mid r_j < \pi_j \qquad \text{Eq.(4)}$$

Yield-Oriented Fitness Function Design

The fitness function is to choose a feature subset that is reasonably compact, agronomically significant, and enhances yield prediction. The validation error of a regression learner trained on a sample set is displayed as prediction error. The pairwise dependence of the chosen variables determines redundancy, while the percentage of chosen variables among all potential variables indicates sparsity. This design avoids having a lot of comparable vegetation or climatic lag factors that don't give the algorithm much more yield information.

$$E(S) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i(S)}{\bar{y}_{r(i)} + \varepsilon}\right)^2 \cdot (1 + \lambda\delta_i)} \qquad \text{Eq.(5)}$$

The prediction error term gives the strongest pressure in the objective. In validation experiments, feature subsets with lower error are favored only when their reduction in error is not obtained by selecting excessive redundant variables.

$$R(S) = 2\sum \left|\mathrm{corr}\left(x_a, x_\beta\right)\right| / [|S|(|S| - 1)] \qquad \text{Eq.(6)}$$

Highly correlated predictors are penalized by a redundancy term. NDVI, EVI and canopy reflectance variables can all be used to show the health of vegetation, however using all of them may impair interpretability and increase model instability. The redundancy penalty pushes the algorithm to preserve the most yield-sensitive representative feature.

Agronomic consistency score indicates if the chosen variables are in accordance with the field of application and is determined based on the average domain weight of the selected features. Features like soil moisture, cumulative temperature, nitrogen availability, vegetation growth and irrigation intensity are given larger prior weights when they are also statistically significant according to the data.

$$F(S) = \omega_e E(S) + \omega_r R(S) + \omega_s |S|/d - \omega_a A(S) \qquad \text{Eq.(7)}$$

Quantum Rotation Updating and Subset Optimization

To update the probability, assess each candidate subgroup and use the best one. The quantum rotation-based feature subset optimization procedure is shown in Figure 2. Rotate feature amplitudes toward inclusion or exclusion after comparing each sampled subset to the best subset as of right now. In this manner, variety is preserved in the early phases and the likelihood of characteristics that consistently enhance yield prediction is progressively increased.

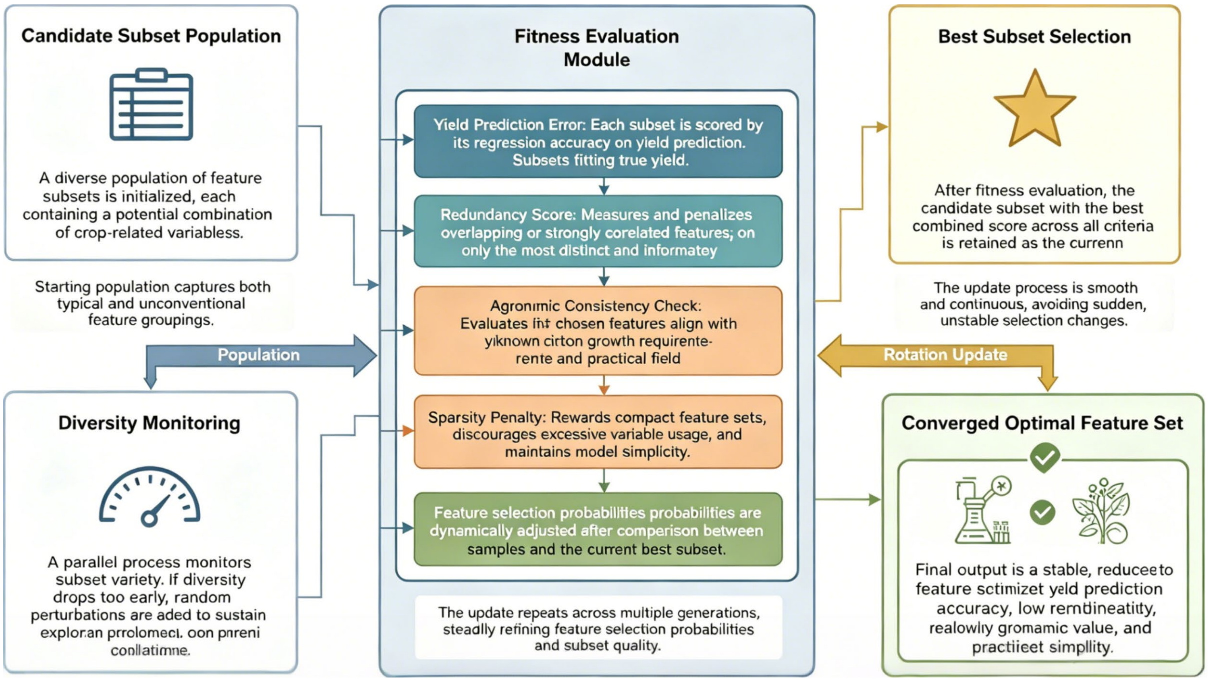

Figure 2. Quantum rotation-based feature subset optimization process

$$\theta_j = \eta\left[F(S_r) - F(S_\beta)\right]h_j \qquad \text{Eq.(8)}$$

Rotate by an amount proportional to the difference in fitness and in the direction indicated by the feature-direction indication. When a sampled subset performs worse than the best subset, the characteristics producing redundancy or weak prediction gain are shifted towards exclusion. Set the learning coefficient to 0.035 for the first 80 iterations and then drop it to 0.018 to ensure convergence.

$$q_j^{\mathrm{new}} = U(\theta_j)q_j \qquad \text{Eq.(9)}$$

The rotation matrix preserves the normalization requirement while altering the probability amplitudes. Instead of abruptly changing a feature state, this method continuously modifies the selection probability and is more stable than direct bit-flipping.

$$D_t = (1/d)\sum_j 2\pi_j(1-\pi_j) \qquad \text{Eq.(10)}$$

The diversity coefficient monitors whether the population collapses too early. If the diversity value falls below 0.12 before the 100th iteration, a small disturbance is added to low-confidence features so that the search can still explore alternative soil-climate combinations.

$$S *= \mathrm{argmin}_s F(S) \qquad \text{Eq.(11)}$$

The final selected subset is the candidate subset with the minimum yield-oriented fitness value. In the default configuration, the algorithm converges within 180 iterations and selects 49 variables from the original 86 - feature pool, producing a 42.5% dimensionality reduction while retaining the main yield-sensitive indicators.

Experimental Configuration and Evaluation Metrics

Dataset Construction and Preprocessing Strategy

The experimental dataset is constructed from multi-source crop production records covering three growing seasons and four agricultural regions. Each sample includes soil nutrient measurements, daily meteorological statistics, remote sensing vegetation indices, irrigation and fertilization records, and management descriptors. The final dataset contains 12480 field-season samples after removing records with incomplete yield labels. The original candidate feature pool contains 86 variables. Missing values below 8% are imputed by region-specific median values, while variables with missing rates above 25% are removed before feature selection. Continuous variables are normalized to the same scale, and temporal weather indicators are aggregated over sowing, vegetative, flowering, grain-filling, and maturity stages.

$$xt_{ij} = \left(x_{ij} - m_j\right)/\left(u_j - l_j + \varepsilon\right) \qquad \text{Eq.(12)}$$

The normalization formula maps heterogeneous agricultural indicators into comparable ranges. This step is necessary because soil potassium, rainfall, temperature accumulation, and vegetation indices have different units and numerical scales. The dataset is divided into 70% training samples, 15% validation samples, and 15% test samples. Cross-region evaluation is additionally performed by training on three regions and testing on the remaining region.

Evaluation Metrics and Comparative Protocol

The four are stability, speed, feature reduction, and prediction accuracy. All-feature modelling, correlation filtering, mutual-information selection, recursive feature removal, LASSO selection, random forest importance selection, genetic algorithm selection and particle swarm feature selection are utilized as comparisons for the suggested method. All of the selected subsets are analyzed using the same random forest regression and gradient boosting regression backbones. Run each procedure 20 times with different random seeds to prevent the problem of random bias. This subsection also contains the comparison protocol, as a standalone protocol-only subsection would lack significant technical material.

$$\mathcal{E}_y(S) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i(S)}{\bar{y} + \varepsilon}\right)^2} \qquad \text{Eq.(13)}$$

Root means square error evaluates prediction deviation in yield units. The proposed method obtains an RMSE of 0.412 tons per hectare on the test set, while all-feature modeling records 0.456 and genetic algorithm selection records 0.439. The improvement indicates that redundant agricultural variables can harm prediction even when a powerful regression model is used.

$$RR = (d - |S|)/d \times 100 \qquad \text{Eq.(14)}$$

The reduction ratio measures how much dimensionality is removed by feature selection. The proposed method reduces the 86 original variables to 49 selected variables, corresponding to a reduction ratio of 42.5%. Correlation filtering removes more variables, but its prediction accuracy is lower because it may discard variables that are individually weak but jointly useful.

Feature stability measures the overlap between subsets selected in repeated runs or different regions. The proposed method reaches an average stability score of 0.781, compared with 0.693 for genetic algorithm selection and 0.658 for particle swarm selection. This difference shows that quantum probability updating produces more stable feature choices under seasonal and regional variation.

Result Interpretation and Comparative Analysis

Yield Prediction Accuracy Analysis

The prediction outcomes of all feature selection techniques are shown in Figure 3. The suggested approach achieves 0.412 tons per hectare, which is less than all-feature modeling at 0.456, mutual-information selection at 0.448, recursive feature elimination at 0.431, and genetic algorithm selection at 0.439 [31]. Figure 3(a) displays the RMSE when employing random forest regression. The coefficient of determination is shown in Figure 3(b), and the suggested approach achieves 0.892, which is 0.047 greater than the all-feature baseline [32]. The distribution of absolute prediction errors is shown in Figure 3(c); the suggested approach has decreased the percentage of samples with errors larger than 0.8 tons per hectare from 13.6% to 8.4% [33]. Feature reduction and the removal of unstable indicators that remain constant throughout the yield's growing stage are the reasons for the improvement.

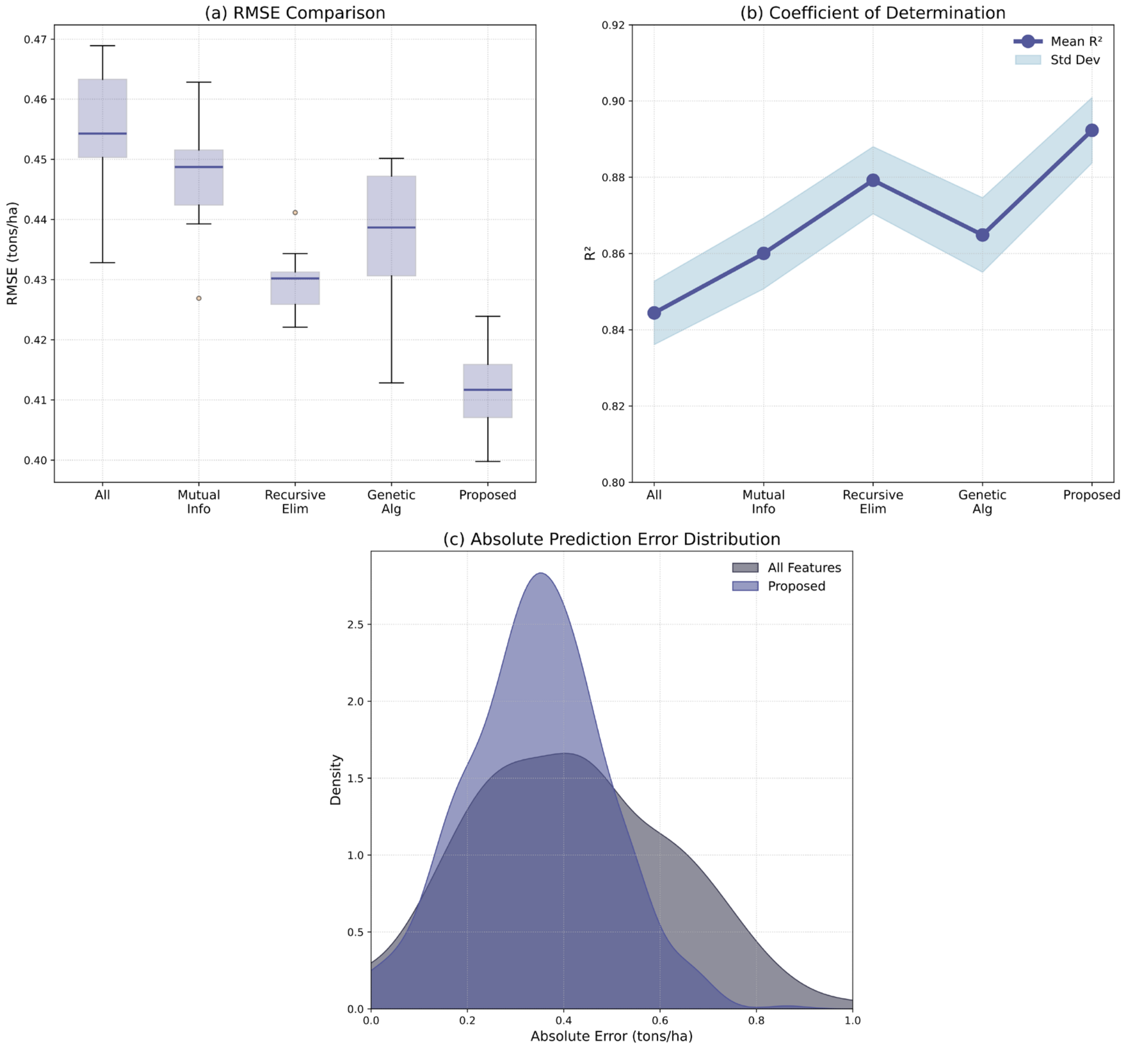

Figure 3. Yield prediction accuracy comparison across feature selection methods. (a) RMSE comparison under regression backbones. (b) Coefficient of determination comparison. (c) Absolute prediction error distribution.

Feature Reduction and Interpretability Analysis

Feature subset redundancy and compactness are depicted in Figure 4. The suggested approach chooses 49 of the 86 variables, as seen in Figure 4(a); recursively removed features choose 56, and genetic algorithm selection chooses 53 [34]. The average pairwise correlation of the chosen variables is displayed in Figure 4(b); the suggested subset has a redundancy score of 0.218, which is less than 0.286 for genetic selection and 0.302 for random forest importance [35]. Because there are less redundant subset assessments following quantum probability convergence, the suggested approach reduces the model training time by 27.3% when compared to wrapper selection, as shown in Figure 4(c) [36]. As a result, the method retains yield-improving information while removing superfluous qualities rather than arbitrarily reducing the number of features.

The distribution of agronomic importance for the chosen traits is displayed in Figure 5. NDVI, EVI, red-edge reflectance, and canopy water indicators are among the remote sensing variables that make up 28.6% of the chosen subset, as shown in Figure 5(a) [37]. The percentage of meteorological factors is 30.6%, as seen in Figure 5(b). The most often chosen meteorological factors are growing-degree days, cumulative rainfall, and flowering-stage temperature variance. Soil organic matter, accessible nitrogen, irrigation frequency, and planting density are all the same in various locations, according to Figure 5(c), which compares soil and management characteristics [38]. The chosen feature structure does not rely on a single data source and is appropriate for displaying crop development characteristics while maintaining information on canopy health, heat accumulation, water supply, nutrient delivery, etc.

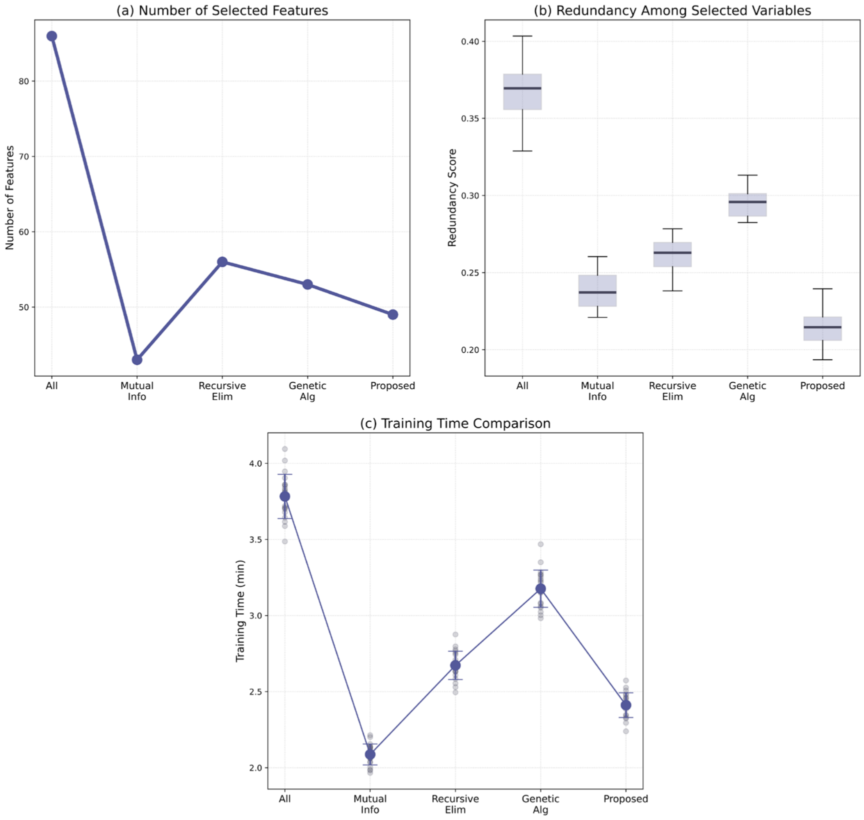

Figure 4. Feature subset compactness and redundancy comparison. (a) Number of selected features. (b) Redundancy score among selected variables. (c) Training time comparison under selected subsets.

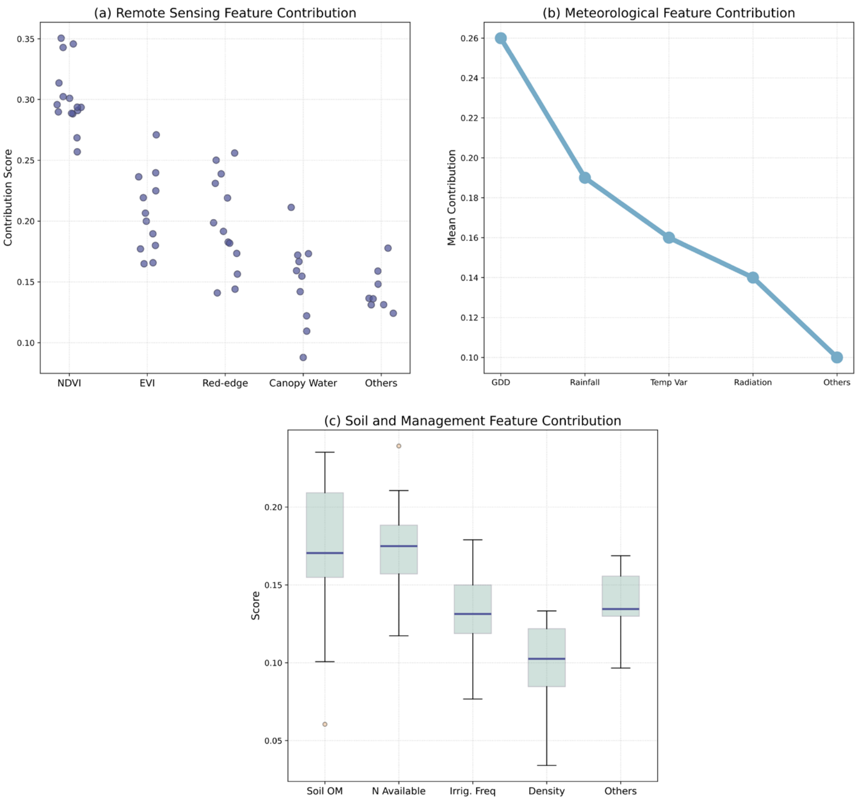

Figure 5. Agronomic importance distribution of selected features. (a) Remote sensing feature contribution. (b) Meteorological feature contribution. (c) Soil and management feature contribution.

Robustness and Generalization Across Crop Scenarios

The outcomes of incomplete and noisy data are displayed in Figure 6. The change in RMSE following the addition of Gaussian noise to meteorological variables is depicted in Figure 6(a); at the greatest noise level, genetic algorithm selection rises to 0.489 and the suggested approach is 0.451 higher than the initial 0.412 [39]. The results of performance at missing rates of 5% to 20% are displayed in Figure 6(b). Because the suggested method chooses a set of various complimentary indicators rather than a small group of strongly correlated characteristics, it still has a reduced error. The comparison of feature stability under noise is shown in Figure 6(c), and even at a 20% missing rate, the suggested approach maintains a stability score higher than 0.72. Therefore, fragile variables that are only useful in clean data are not required for the quantum rotation update.

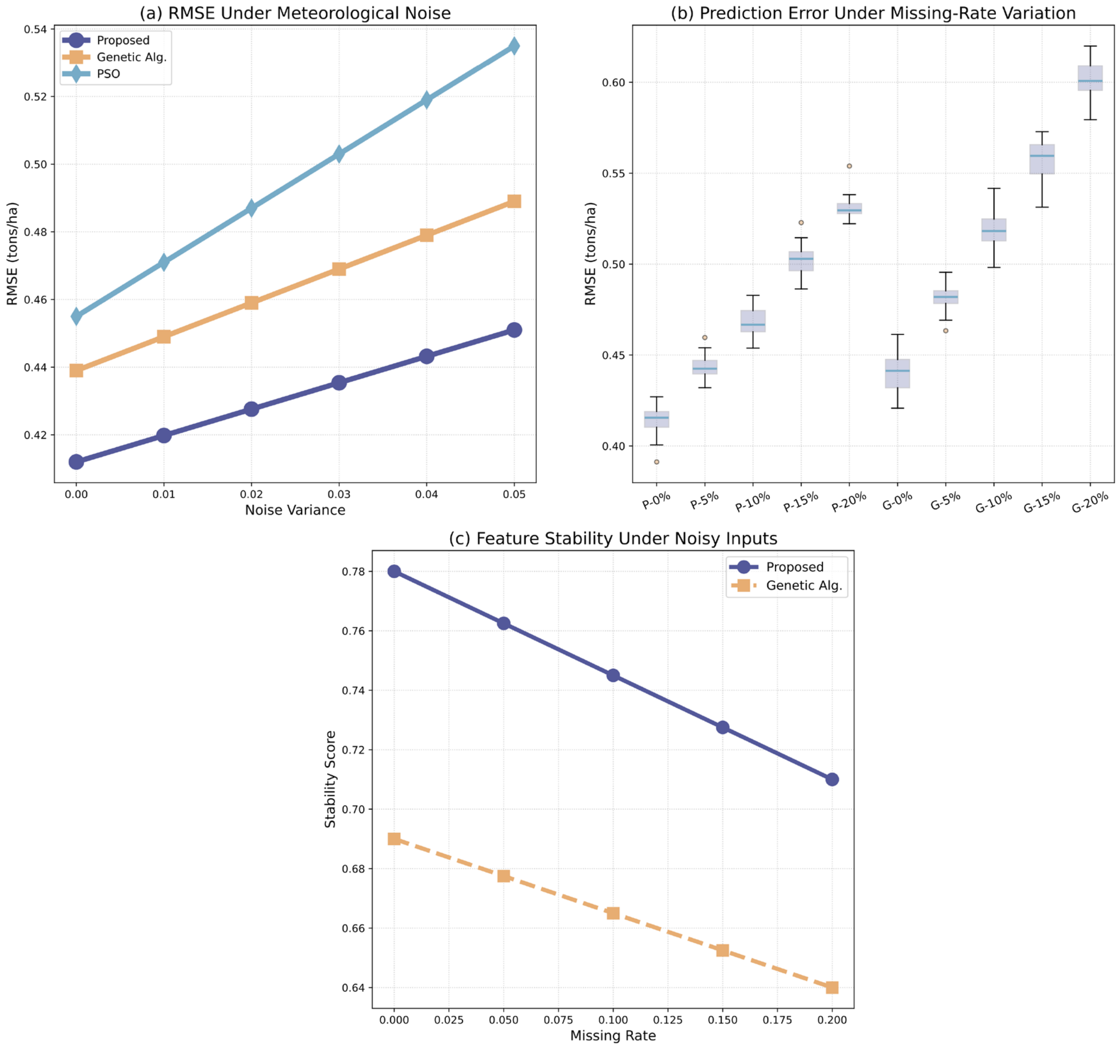

Figure 6. Robustness analysis under different data noise and missing-rate conditions. (a) RMSE under meteorological noise. (b) Prediction error under missing-rate variation. (c) Feature stability under noisy inputs.

Cross-season and cross-region generalization are displayed in Figure 7. The suggested approach outperforms the all-feature model at 0.521 tons/ha and genetic algorithm selection at 0.497 tons/ha, as seen in Figure 7(a), with an average regional transfer RMSE of 0.463 tons/ha [40]. The cross-season prediction accuracy is displayed in Figure 7(b), which illustrates that the chosen features remain useful even after variations in the distribution of heat accumulation and rainfall. A comparison of feature overlap by season is shown in Figure 7(c). The suggested approach maintained an overlap score of 0.755, suggesting that it might identify yield drivers that are comparatively stable. Because redundancy suppression keeps the predictor from overfitting to season-specific weather noise, the greatest generalization benefit happens in regions with substantial climate variability.

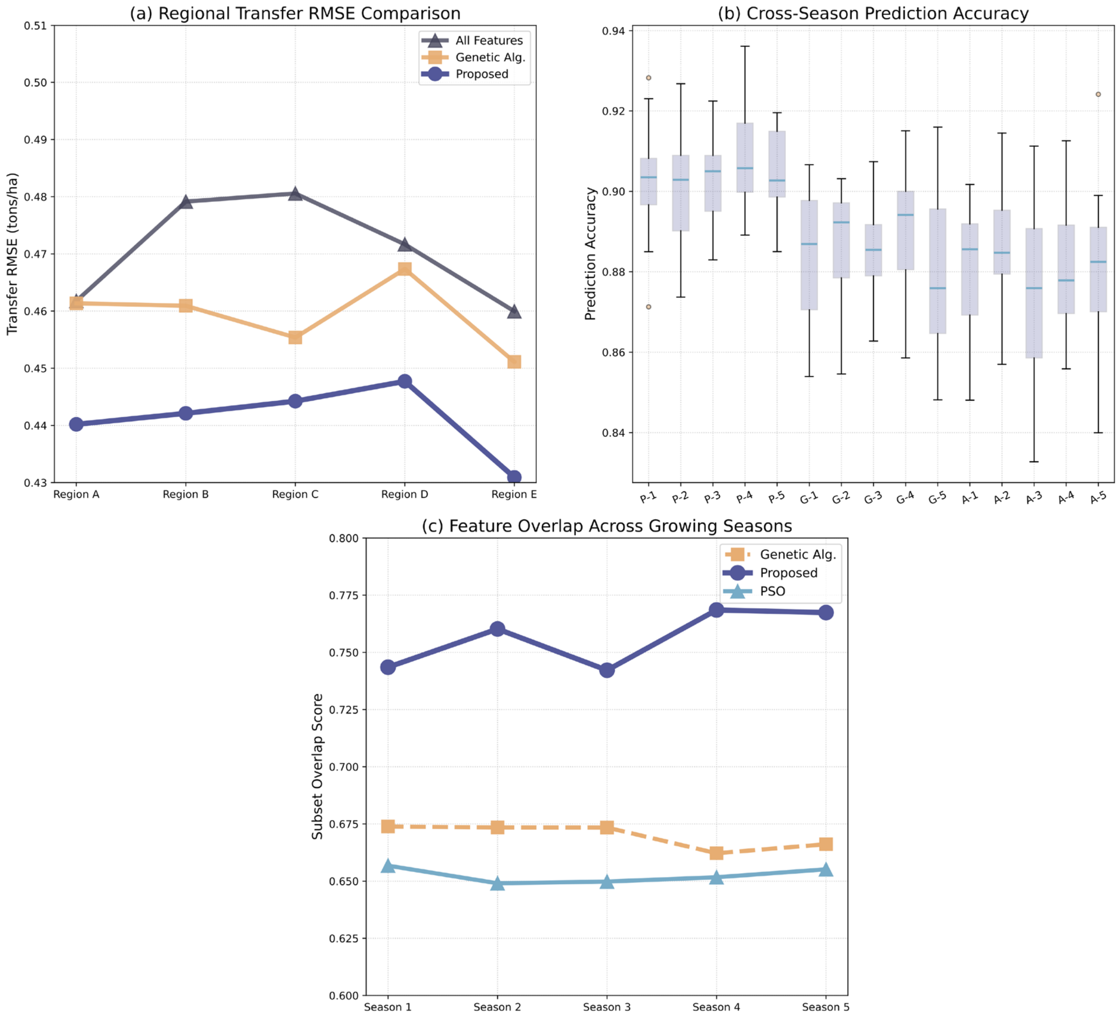

Figure 7. Cross-region and cross-season generalization performance. (a) Regional transfer RMSE comparison. (b) Cross-season prediction accuracy. (c) Feature overlap across growing seasons.

Conclusion

By eliminating unnecessary agricultural variables, quantum-assisted feature selection has improved crop production optimization without compromising prediction accuracy. The approach uses quantum rotation to update the selection probability, evaluates subsets using a yield-oriented fitness function, and suggests candidate variables as probability amplitudes. Compared to conventional feature selection techniques, it is more agronomically interpretable, computationally efficient, and offers a rather fair trade-off between prediction accuracy and dimension reduction.

The following are the initial findings of the study. First, there is a lot of redundancy in the high-dimensional agricultural data, such as between correlated soil indicators, vegetation indices, and weather-lag factors. Second, the problem of early convergence is lessened and the search is more diverse thanks to quantum probability representation. Third, a more stable feature subset is produced by taking into account prediction error, redundancy penalty, sparsity control, and agronomic consistency rather than just accuracy. Fourth, the chosen variables are appropriate for actual yield management and have demonstrated greater cross-region and cross-season stability.

In the future, online agricultural sensor networks, crop growth simulation models, and actual quantum annealing platforms could be integrated with quantum-assisted selection. More work will be done in the future to identify features that are appropriate for all crops, soil types, and climate-risk places. Quantum-assisted feature optimization cannot be used in intelligent precision agricultural systems without improved data quality and hardware support.

Author Contributions

Darius Cojocaru contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, project administration, and funding acquisition. Sandu Oancea contributes to conceptualization, methodology, software, validation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. Frontiers in Plant Science, 10, 621. https://doi.org/10.3389/fpls.2019.00621

[2] Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture, 163, 104859. https://doi.org/10.1016/j.compag.2019.104859

[3] van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709

[4] Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific Reports, 11, 1606. https://doi.org/10.1038/s41598-020-80820-1

[5] Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., & Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. Agricultural Systems, 187, 103016. https://doi.org/10.1016/j.agsy.2020.103016

[6] Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agricultural and Forest Meteorology, 274, 144-159. https://doi.org/10.1016/j.agrformet.2019.03.010

[7] Wang, A. X., Tran, C., Desai, N., Lobell, D., & Ermon, S. (2019). Deep transfer learning for crop yield prediction with remote sensing data. Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, 50, 1-5. https://doi.org/10.1145/3314344.3332482

[8] Kamilaris, A., & Prenafeta-Boldu, F. X. (2019). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90. https://doi.org/10.1016/j.compag.2018.02.016

[9] Schuld, M., & Killoran, N. (2019). Quantum machine learning in feature Hilbert spaces. Physical Review Letters, 122, 040504. https://doi.org/10.1103/PhysRevLett.122.040504

[10] Havlicek, V., Corcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. Nature, 567, 209-212. https://doi.org/10.1038/s41586-019-0980-2

[11] Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., & Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. Remote Sensing of Environment, 237, 111599. https://doi.org/10.1016/j.rse.2019.111599

[12] Shook, J., Gangopadhyay, T., Wu, L., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2021). Crop yield prediction integrating genotype and weather variables using deep learning. PLOS ONE, 16, e0252402. https://doi.org/10.1371/journal.pone.0252402

[13] Miao, Y., Mulla, D. J., Robert, P. C., & Hernandez, J. A. (2020). Within-field corn yield prediction using hyperspectral remote sensing imagery. Precision Agriculture, 21, 102-122. https://doi.org/10.1007/s11119-019-09654-5

[14] Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., & Peng, J. (2019). Toward building a transparent statistical model for improving crop yield prediction. Field Crops Research, 234, 55-65. https://doi.org/10.1016/j.fcr.2019.02.005

[15] Kuwata, K., & Shibasaki, R. (2020). Estimating crop yields with deep learning and remotely sensed data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 492-500. https://doi.org/10.1109/JSTARS.2020.2965126

[16] Elavarasan, D., & Vincent, P. M. D. (2020). Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. IEEE Access, 8, 86886-86901. https://doi.org/10.1109/ACCESS.2020.2992480

[17] Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2019). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture. Precision Agriculture, 19, 895-921. https://doi.org/10.1007/s11119-018-9578-7

[18] Balducci, F., Impedovo, D., & Pirlo, G. (2019). Machine learning applications on agricultural datasets for smart farm enhancement. Machines, 6, 38. https://doi.org/10.3390/machines6030038

[19] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2019). Machine learning in agriculture: A review. Sensors, 18, 2674. https://doi.org/10.3390/s18082674

[20] García, S., Luengo, J., & Herrera, F. (2020). Data preprocessing in data mining. Intelligent Systems Reference Library, 72, 1-320. https://doi.org/10.1007/978-3-319-10247-4

[21] Bolon-Canedo, V., Sanchez-Marono, N., & Alonso-Betanzos, A. (2019). Recent advances and emerging challenges of feature selection in the context of big data. Knowledge-Based Systems, 86, 33-45. https://doi.org/10.1016/j.knosys.2015.05.014

[22] Cai, J., Luo, J., Wang, S., & Yang, S. (2019). Feature selection in machine learning: A new perspective. Neurocomputing, 300, 70-79. https://doi.org/10.1016/j.neucom.2017.11.077

[23] Zhang, Y., Gong, D. W., Sun, X. Y., & Guo, Y. N. (2019). A PSO-based multi-objective multi-label feature selection method in classification. Scientific Reports, 7, 376. https://doi.org/10.1038/s41598-017-00416-0

[24] Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2019). A survey on evolutionary computation approaches to feature selection. IEEE Transactions on Evolutionary Computation, 20, 606-626. https://doi.org/10.1109/TEVC.2015.2504420

[25] Zhang, Y., Zhou, Y., Zhang, K., & Pan, Q. (2019). Quantum-inspired evolutionary algorithm for feature subset selection. Applied Soft Computing, 82, 105542. https://doi.org/10.1016/j.asoc.2019.105542

[26] Wang, H., Yang, J., & Liu, H. (2020). Quantum-behaved particle swarm optimization for feature selection. Expert Systems with Applications, 152, 113377. https://doi.org/10.1016/j.eswa.2020.113377

[27] Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. (2019). Quantum machine learning. Nature, 549, 195-202. https://doi.org/10.1038/nature23474

[28] Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., et al. (2021). Variational quantum algorithms. Nature Reviews Physics, 3, 625-644. https://doi.org/10.1038/s42254-021-00348-9

[29] Moll, N., Barkoutsos, P., Bishop, L. S., Chow, J. M., Cross, A., Egger, D. J., et al. (2019). Quantum optimization using variational algorithms on near-term quantum devices. Quantum Science and Technology, 3, 030503. https://doi.org/10.1088/2058-9565/aab822

[30] McClean, J. R., Romero, J., Babbush, R., & Aspuru-Guzik, A. (2019). The theory of variational hybrid quantum-classical algorithms. New Journal of Physics, 18, 023023. https://doi.org/10.1088/1367-2630/18/2/023023

[31] Perdomo-Ortiz, A., Benedetti, M., Realpe-Gomez, J., & Biswas, R. (2019). Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers. Quantum Science and Technology, 3, 030502. https://doi.org/10.1088/2058-9565/aab859

[32] Dunjko, V., & Briegel, H. J. (2019). Machine learning and artificial intelligence in the quantum domain. Reports on Progress in Physics, 81, 074001. https://doi.org/10.1088/1361-6633/aab406

[33] Benedetti, M., Lloyd, E., Sack, S., & Fiorentini, M. (2019). Parameterized quantum circuits as machine learning models. Quantum Science and Technology, 4, 043001. https://doi.org/10.1088/2058-9565/ab4eb5

[34] Farhi, E., Goldstone, J., & Gutmann, S. (2020). A quantum approximate optimization algorithm. MIT Research Reports, 1, 1-16. https://doi.org/10.48550/arXiv.1411.4028

[35] Guerreschi, G. G., & Matsuura, A. Y. (2019). QAOA for Max-Cut requires hundreds of qubits for quantum speed-up. Scientific Reports, 9, 6903. https://doi.org/10.1038/s41598-019-43176-9

[36] Mücke, S., Heese, R., Müller, S., Wolter, M., & Piatkowski, N. (2022). Feature selection on quantum computers. Quantum Machine Intelligence, 5, 11. https://doi.org/10.1007/s42484-023-00096-y
[37] Pudenz, K. L., Albash, T., & Lidar, D. A. (2019). Quantum annealing correction for random Ising problems. Physical Review A, 91, 042302. https://doi.org/10.1103/PhysRevA.91.042302
[38] Lucas, A. (2019). Ising formulations of many NP problems. Frontiers in Physics, 2, 5. https://doi.org/10.3389/fphy.2014.00005
[39] Niu, M. Y., Boixo, S., Smelyanskiy, V. N., & Neven, H. (2019). Universal quantum control through deep reinforcement learning. npj Quantum Information, 5, 33. https://doi.org/10.1038/s41534-019-0141-3
[40] Gidney, C., & Ekerå, M. (2021). How to factor 2048-bit RSA integers in 8 hours using 20 million noisy qubits. Quantum, 5, 433. https://doi.org/10.22331/q-2021-04-15-433