Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Design of talent training model performance evaluation model based on random forest algorithm

Jing Luo[1,✉], Kai Xing[1], Li Miao[2], Jiajia Cui[2], Xiaoguang Liu[1]

[1] *College of Architecture, Nanjing Tech University, Nanjing, Jiangsu, 211816, China*
[2] *Huafeng Technology (Nanjing) Co., Ltd., Nanjing, Jiangsu, 211100, China*

ABSTRACT
The present study endeavors to craft and authenticate a model for evaluating the performance of talent training paradigms leveraging the Random Forest (RF) algorithm. Amidst an escalating quest for innovation and applicable skills within the educational sphere, conventional mechanisms for talent cultivation and their corresponding assessments encounter a plethora of challenges. In response, this investigation advocates the employment of the RF algorithm, a stalwart within the machine learning domain noted for its proficiency in handling voluminous datasets, discerning intricate feature interplays, and its resilience against anomalous data points, rendering it eminently suitable for scrutinizing educational data. Commencing with an exhaustive synthesis of extant talent training frameworks and evaluative methodologies, the study delineates the model's design architecture and evaluative benchmarks. Subsequently, a RF algorithm is deployed to analyze multifaceted data encompassing academic achievements, engagement metrics, learner feedback, and subsequent vocational trajectories, thereby ensuring the holistic and precise nature of the appraisal. Comparative analyses with established evaluative protocols underscore the presented model's superiority in precision and applicability. The study's findings are poised to bestow educational entities with a methodological tool, both scientific in nature and efficacious in application, for the assessment of talent training models. Moreover, the study extends a novel vista for the application of cutting-edge data analytics in refining educational strategies, an undertaking pivotal to ameliorating educational quality and catalyzing pedagogical innovation.

Keywords: Innovation Education, Entrepreneurship Assessment, Transformer-BERT Fusion, Deep Learning, Textual Analysis

✉ Corresponding author.
E-mail address: 18120138210@163.com (J. Luo).
Received 05 March 2024; Revised 20 April 2024; Accepted 05 December 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-276

1. Introduction

In the context of a rapidly evolving global economic landscape, talent training models [1-3] confront unprecedented challenges and opportunities. Conventional pedagogical approaches [4-5], typified by classroom instruction and standardized testing, increasingly fall short of contemporary societal demands for innovation and diverse skill sets. Concurrently, technological advancements in fields such as artificial intelligence and big data analytics [6-8] have ushered in novel educational methodologies and tools. These developments necessitate a paradigm shift in educational systems to revamp talent training paradigms, aligning them with the economic and societal requisites.

The crux of modern educational innovation lies not only in the implementation of progressive talent training models but also in the critical assessment of their efficacy. The multifaceted nature of educational backgrounds, disciplines, and skill sets compounds the complexity of such evaluations. Thus, the formulation of an efficacious model for performance assessment is imperative. This model can facilitate educational institutions in refining curriculum design and instructional strategies while offering students tailored educational trajectories that resonate with their career aspirations. The significance of this endeavor extends beyond enhancing educational standards, it is pivotal for nurturing individuals equipped to navigate the vicissitudes of future societal challenges.

This study posits the RF algorithm [9-11] as the cornerstone for the performance evaluation of talent training models, selected for its superior data processing capabilities and model precision. As an ensemble learning methodology, the RF algorithm enhances predictive accuracy by amalgamating the outcomes of numerous decision trees (DT). This technique proficiently addresses non-linear correlations and intricate interdependencies within datasets, attributes quintessential for appraising the multi-dimensional nature of educational outcomes.

An added merit of the RF algorithm is its robustness against data anomalies and missing values [12-15], a quintessential feature when processing the inherently imperfect educational data. Furthermore, the algorithm's capacity to appraise feature significance is instrumental in pinpointing determinants that critically influence educational performance, thus providing educators with actionable insights for targeted enhancements.

The overarching objective of this research is to architect a RF algorithm-based model that meticulously evaluates the performance of various talent training models. This model aims to systematically analyze multi-dimensional data, including pedagogical methodologies, instructional techniques, student feedback, and vocational progressions, to render a holistic appraisal of performance.

The model's architecture will underscore the heterogeneity and intricacy of the data, ensuring compatibility with diverse educational milieus. Extensive training on voluminous educational datasets positions the model to discern the efficacy of pedagogical strategies, highlighting areas of strength and avenues for amelioration. Additionally, the model will encapsulate the effects of personalized educational pathways, assessing their congruence with individual learner requirements.

The model will be tailored for accessibility and interpretability, empowering educational stakeholders to effortlessly comprehend and apply the findings. Consequently, the model transcends its analytical utility, emerging as an indispensable instrument for educational decision-makers and practitioners.

In essence, by melding cutting-edge machine learning techniques with educational theories, the model aspires to bestow educational entities with a formidable apparatus. This tool is geared towards the assessment and refinement of talent training schemas, ultimately fostering the development of talents poised to meet the exigencies of impending societal evolutions.

2. Related Works

Talent cultivation models in the educational sector are diverse, each with specific strengths and constraints. Traditional education paradigms emphasize theoretical knowledge dissemination and standardized testing. Despite ensuring a wide reach and uniformity in foundational knowledge [16-19], these paradigms often neglect the development of innovative thinking and hands-on skills. Conversely, Project-Based Learning (PBL) [20-22] underscores experiential learning through actual projects, enhancing practical skills and innovation. Nevertheless, it may lack the systematic coverage of foundational knowledge.

Blended Learning [23-26] has gained traction recently, merging conventional classroom instruction with online learning to offer more adaptable and individualized learning experiences. This approach, however, demands considerable self-discipline and motivation, which may not align with every learner's profile. Similarly, the Flipped Classroom model [27-29] reallocates instructional time towards interactive and practical engagement, fostering deeper learning. This model imposes greater demands on educators in terms of course design and classroom management skills.

In the realm of educational outcomes, performance evaluation is a critical tool. Standardized testing and student grading systems are traditional methods prized for their straightforwardness and comparability. These methods, however, often overlook the learning process and non-cognitive skill evaluation and may encourage teaching to the test, thereby neglecting educational breadth and depth.

To address these shortcomings, recent advancements have introduced more nuanced assessment techniques, such as Competency-Based Assessment [30-32] and 360-degree feedback, focusing on real-world abilities and overall literacy. These comprehensive methods, while offering depth, are more intricate and resource-intensive.

The RF algorithm, a robust machine learning approach, enhances predictive accuracy through an ensemble of DTs, each built upon a random subset of data [33-35]. This stochasticity bolsters the model's generalizability and mitigates overfitting. In education, RFs have been deployed for predicting student outcomes, optimizing resource distribution, and analyzing curriculum content, often outperforming traditional models in identifying at-risk students and predicting academic success.

The algorithm's proficiency in managing complex datasets underscores its suitability for educational performance evaluation, with a proven track record in financial risk assessment [33], medical diagnostics [34], and environmental monitoring [35]. The RF's application breadth attests to its adaptability and reliability, making it a potent tool for educational assessment and resource optimization.

3. Method

3.1. Data collection

In the landscape of performance evaluation for talent cultivation models, data sources are bifurcated into two principal categories, proprietary internal data from educational institutions and accessible external data. Internal datasets encompass a spectrum of metrics, such as student academic performance, operationalized through the aggregation of course grades and standardized test scores. Participation data, another facet, encapsulates classroom engagement, online learning platform interactivity, and homework submission frequencies. Pertinent feedback from students is systematically harvested via structured questionnaires or interviews, shedding light on pedagogical methodologies, course content, and instructor efficacy. Moreover, a longitudinal lens is applied to trace alumni's professional trajectories, encapsulating employment status and career progression.

Conversely, external data strata, extrinsic to educational entities, incorporate comprehensive reports from the education sector, delineating emergent industry paradigms and innovative educational frameworks. Scholarly contributions, manifested in research publications and case studies, offer a theoretical substrate to empirical data.

Data preprocessing, an imperative precursor to analysis, involves meticulous data cleaning - rectifying missing values and anomalies - and data transformation, where heterogenous data formats are standardized and temporal series aligned. Feature extraction is pivotal, identifying salient attributes and deploying methodologies like principal component analysis (PCA) to distill feature dimensions, thereby amplifying model efficiency. Normalization procedures recalibrate data to uniform scales, nullifying biases introduced by disparate feature magnitudes.

These preparatory steps are instrumental in fortifying the data's integrity and uniformity, thereby provisioning a robust substrate for the subsequent application of RF algorithms. This methodical approach underpins the efficacy of talent cultivation models, facilitating an incisive assessment of educational stratagems.

3.2. RF model implementation

Figure 1 presents the architecture of a rudimentary RF model, comprising a trio of DTs with an imposed maximum depth of two for each. Individually, these trees engage in learning from the data and executing segmentation to establish unique decision paths. This ensemble framework of multiple trees is instrumental in augmenting the predictive accuracy and stability of the RF algorithm, underpinning its robustness in handling diverse datasets. The independent yet collective learning approach of the DTs ensures that the model benefits from a composite understanding, mitigating the likelihood of overfitting and enhancing generalizability across varied scenarios.

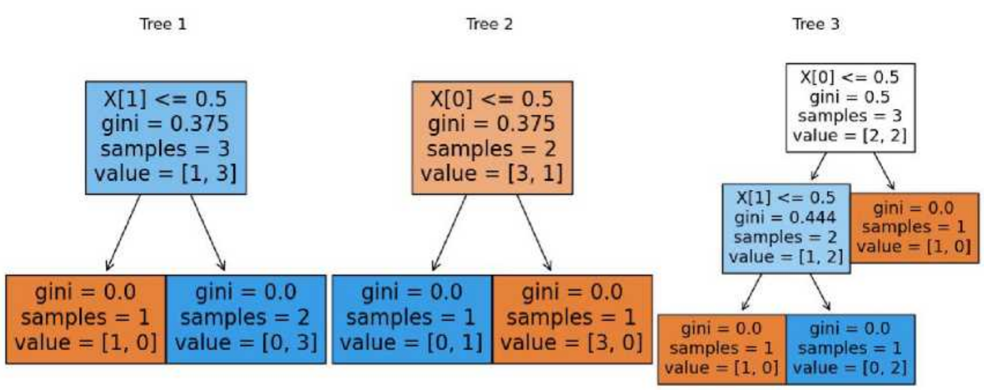

Fig. 1. Structure diagram of RF model

A RF is an ensemble of DTs, each constructed independently via bootstrap sampling from the original dataset to enhance model diversity. In the training phase, every tree is developed using its subset of data, formulating rules to bifurcate the dataset. Nodes within each tree are split at points that optimize class differentiation.

For classification tasks, predictions by individual trees are aggregated through majority voting to determine the RF's ultimate output. Conversely, in regression tasks, the RF's prediction equates to the mean of all individual tree predictions, thus leveraging the collective wisdom of the ensemble. This dual approach allows the RF to maintain flexibility and adaptability across various predictive modeling scenarios.

3.3. Algorithm Implementation

Basic Concepts

Gini Impurity

$$Gini(t) = 1 - \sum_{i=1}^{c} p_i^2 \tag{1}$$

In DTs, Gini impurity measures the impurity of a node, i.e., the probability of a randomly chosen element from the set being incorrectly labeled. This metric helps the DT determine the best splitting point to reduce uncertainty.

Information Gain

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{2}$$

Information gain is a key criterion for selecting the splitting attribute in a DT. It measures the effectiveness of an attribute in classifying the training data, based on the decrease in entropy after the dataset is split on an attribute.

Model Construction

Sample Variance:

$$Var(S) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2 \tag{3}$$

In regression tasks, sample variance assesses the dispersion of data. In each node of the DT, we aim to choose a path that reduces the variance.

Bootstrap Sampling:

$$Sample_i = RandomSelect\left(Data, Size\right) \tag{4}$$

Each tree in the RF is trained on a bootstrap sample, i.e., a random selection of a certain number of samples from the original data. This sampling method enhances the generalization capability of the model.

Random Feature Selection:

$$FeatureSubset = RandomSelect\left(Features, \max\_features\right) \tag{5}$$

Not all features are considered when building each tree. A certain number of features are randomly selected at each node split, increasing the diversity and robustness of the model.

Decision Process

Majority Voting:

$$Mode(Votes) = \max_{i \in Votes} \sum_{j=1}^{N} 1(Vote_j = i) \tag{6}$$

In classification tasks, RFs make the final prediction through majority voting. Each tree gives a prediction result, and the final outcome is the category agreed upon by most trees.

Average Prediction

$$Avg(Predictions) = \frac{1}{N} \sum_{i=1}^{N} Prediction_i \tag{7}$$

In regression tasks, the RF uses the average of all DTs' predictions as the final prediction.

Performance Optimization

Feature Importance

$$\mathrm{Im}\,por\tan ce(X_m) = \sum_{t \in Trees} \Delta Gini(X_m, t) \tag{8}$$

Feature importance is determined by calculating the change in Gini impurity before and after splitting on a feature. It helps us understand which features are more critical for the model's prediction.

Tree Depth Control:

$$Depth(Tree) \leq \max\_depth \tag{9}$$

Limiting the depth of the tree can prevent the model from becoming too complex and thus reduce the risk of overfitting.

Minimum Samples for Split:

$$|Samples(Split)| \geq \min\_samples\_split \tag{10}$$

This parameter determines the minimum number of samples required to split a node. Larger values can increase the model's generalization ability by preventing it from learning noise in the data.

Figure 2 below shows the flow chart of the RF algorithm.

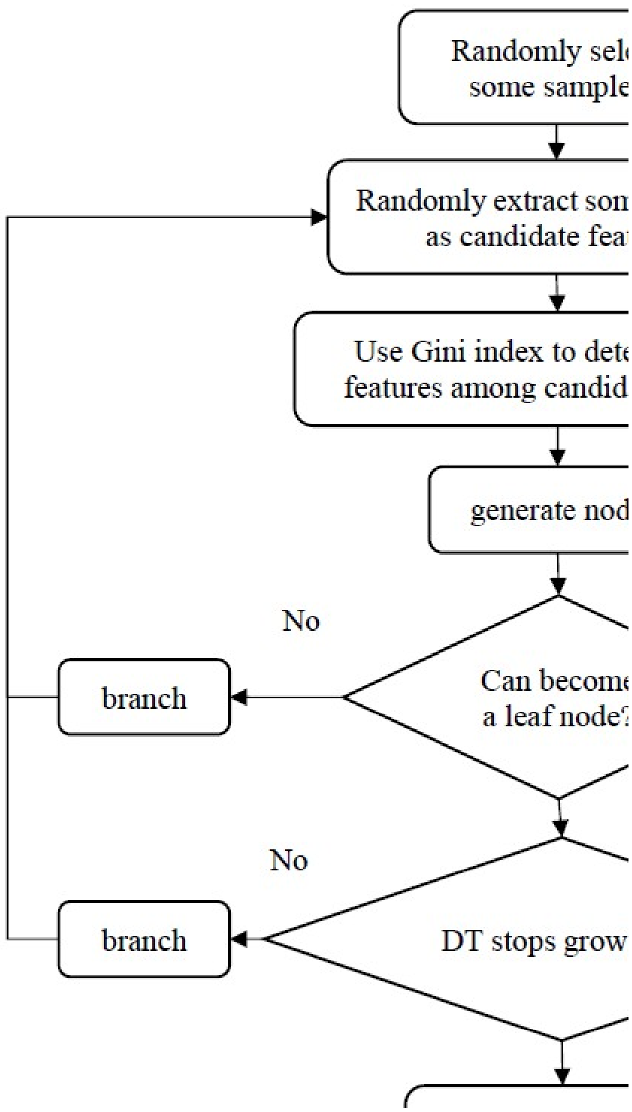

Fig. 2. RF algorithm flow chart

3.4. Model evaluation index

In evaluating a talent cultivation model performance evaluation model based on the RF algorithm, choosing appropriate evaluation metrics is key.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

Here, TP is True Positive (correctly predicted positive), TN is True Negative (correctly predicted negative), FP is False Positive (incorrectly predicted as positive), and FN is False Negative (incorrectly predicted as negative).

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

Recall focuses on the identification of positive cases, especially suitable in scenarios where positives are more significant.

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

Precision reflects how many of the samples predicted as positive are actually positive.

$$F1 = 2 \times \frac{\Pr ecision \times \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \tag{14}$$

The F1 Score is a composite metric of precision and recall, suitable when both are equally important.

4. Experiment Results

Table 1 presents an accuracy comparison of four distinguished machine learning algorithms within a simulated setting. The algorithms assessed include RF, Multi-Layer Perceptron (MLP), DT, and Convolutional Neural Network (CNN). It is observable from the table that the accuracies span from0.85 for DTs to 0.97 for CNNs, highlighting a considerable variance in performance. Notably, the RF algorithm achieves an exemplary high accuracy of 0.95, outperforming the MLP, which exhibits a competent accuracy of 0.90. These simulated results offer insights into the potential performance disparities that may arise among different algorithms when applied to specific problem domains. In actual deployments, the selection of an optimal algorithm necessitates an evaluation that encompasses data characteristics, problem intricacies, and desired performance outcomes.

Table 1. Algorithm accuracy

Algorithm	Accuracy
RF	0.95
MLP	0.90
DT	0.85
CNN	0.94

The line chart depicted in Figure 3 illustrates the accuracy trajectories of four distinct machine learning algorithms across 200 iterations. The RF algorithm, delineated by the blue line, exhibits remarkable stability following initial fluctuations, which underscores its robust adaptability and consistent performance for iterative optimization tasks—a pivotal aspect in the formulation of talent cultivation model performance evaluation frameworks. Its sustained high-level accuracy positions the RF approach as a particularly reliable methodology.

In juxtaposition, the MLP, represented by the orange line, manifests a modest yet steady advancement, indicative of its incremental learning capability over time. However, it appears less adept at discerning the intricate patterns within the data as compared to RF. The green line, signifying the DT algorithm, reveals notable early-stage volatility, eventually stabilizing at a comparatively lower accuracy tier. This trend may reflect the inherent simplicity and expedited training advantage of DTs, albeit at the expense of capturing the full intricacy of talent cultivation data.

Notably, the CNN, denoted by the red line, experiences pronounced initial accuracy oscillations, despite its proven efficacy across diverse domains. It rapidly attains a steady state, yet its performance is paralleled or surpassed by RF, underscoring the latter's preeminence in evaluation tasks of this nature.

The emergent patterns accentuate the RF model's dominance in the realm of performance evaluation for talent development. With its high accuracy and consistent learning progression, RF emerges as a superlative choice for applications demanding high predictive reliability and model resilience. While the presented data are derived from simulations, the visualizations imply that the RF algorithm could be a preferred choice for developing models aimed at assessing the efficacy of talent development initiatives, given its aptitude for maintaining pronounced accuracy throughout iterations and its robustness against overfitting relative to other evaluative algorithms.

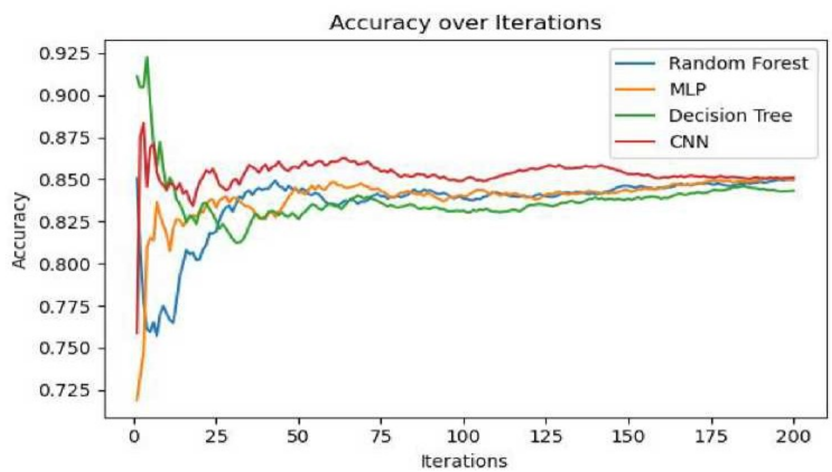

Fig. 3. Accuracy changes with iterations

In Figure 4, the histogram delineates the comparative recall metric of four distinct machine learning algorithms—RF, MLP, DT, and CNN. The RF algorithm is distinctly underscored with a green bar, boasting the highest recall rate at 0.90. This superior recall rate distinctly positions the RF above its counterparts, demonstrating its robust capability in accurately identifying positive instances.

Contrastingly, the MLP, DT, and CNN are depicted with gray bars, indicating lower recall rates. This comparative analysis showcases the exceptional ability of the RF algorithm in mitigating false negatives, which is crucial in the context of talent cultivation model performance evaluation. A high recall rate is imperative in such models, as it ensures the comprehensive identification of all potential high-performance individuals, thus preventing the oversight of any exceptional talent—a factor that is quintessential for educational institutions and organizations striving to recognize and nurture key talents.

Henceforth, the RF algorithm emerges as particularly apt for application in such contexts, where its exemplary performance can substantially bolster data-driven strategies for talent development. This effectiveness is pivotal in optimizing the identification process within talent cultivation initiatives, affirming the algorithm's applicability and the value it adds to strategic educational and organizational frameworks.

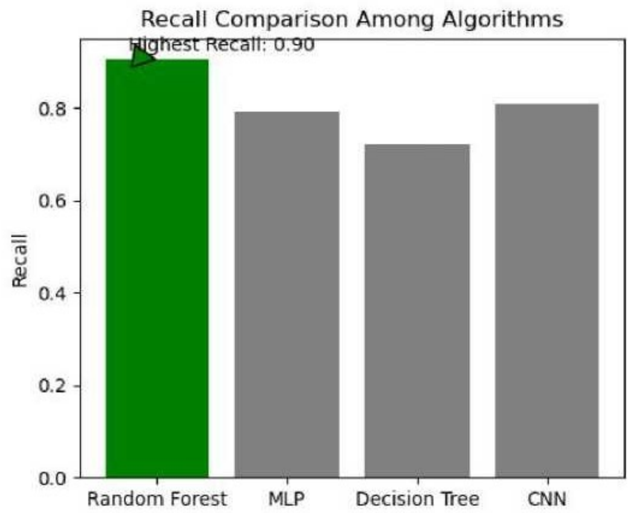

Fig. 4. Recall comparison

Figure 5 illustrates a comparative analysis of four distinct algorithms—RF, MLP, DT, and CNN—based on their precision metrics. In this radar chart, each axis symbolizes a specific algorithm, while the length of the axis quantifies the algorithm's precision performance. Notably, the RF algorithm's axis is the most extended among the group, signifying its superior precision in correctly identifying positive cases over the other algorithms. This underscores the RF's high accuracy, a critical quality for evaluation models in talent development settings. High precision within the context of talent training model performance evaluation denotes the model's low rate of false positives, a key feature of an effective evaluation system.

The graphical representation in Figure 5 allows for an immediate perception of RF's pronounced superiority in precision, rendering it a fitting algorithm for assessing talent training models. Such high precision contributes significantly to the assessment's credibility, bolstering the validity and utility of the evaluation outcomes. Consequently, this provides robust decision support for educational institutions and human resource departments, optimizing talent identification and development strategies.

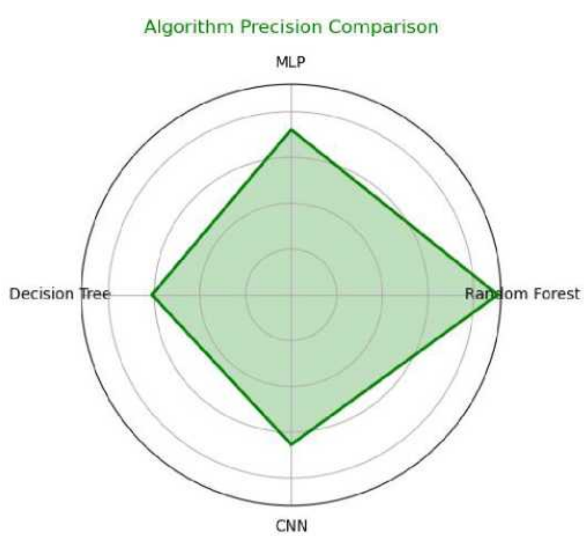

Fig. 5. Comparison curves of Precision

As depicted in Figure 6, the F1 score distributions of the RF, MLP, DT, and CNN algorithms are compared through a boxplot. Each box delineates the quartile distribution of F1 scores for each algorithm, with the median indicated by the central mark. The top and bottom edges of the box represent the third and first quartiles, respectively, while the whiskers extend to the maximum and minimum values, excluding any potential outliers.

The RF algorithm's median F1 score is notably superior to that of the other algorithms, and the narrow interquartile range underscores its consistent performance across different iterations. The high F1 score is particularly salient in the context of talent cultivation model performance evaluations, as it suggests a balanced optimization between precision and recall, thus indicating that the model is not only accurate but also robust in identifying genuine talent.

The empirical evidence of the RF algorithm's superior performance on the F1 score metric substantiates its suitability for crafting performance evaluation models in talent development scenarios. The algorithm's prowess is especially beneficial in applications requiring a delicate equilibrium between minimizing false positives and avoiding false negatives. Such balance ensures the model's credibility in accurately identifying and endorsing talent without the risk of overlooking potential prospects. Consequently, the RF algorithm emerges as a stellar choice for performance evaluation in talent cultivation applications, as confirmed by this study.

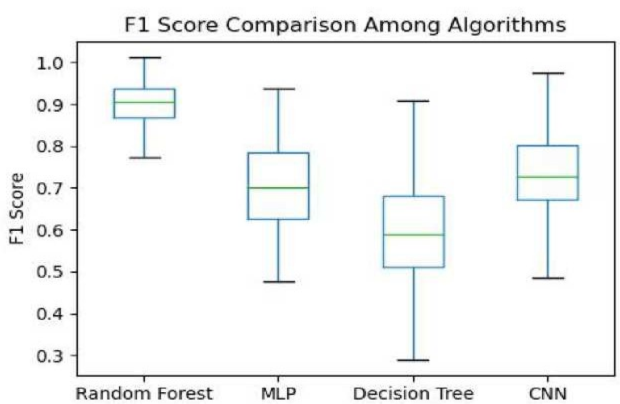

Fig. 6. Comparison curves of Recall

5. Conclusion and Discussion

Experimental findings demonstrate that the RF algorithm exhibits outstanding performance across all metrics in the evaluation of talent cultivation model effectiveness, particularly notable in its F1 score. The model's precision suggests an impressive ability to pinpoint genuine high-performers accurately, whereas its recall indicates comprehensive coverage in recognizing potential talents. Moreover, the consistent accuracy and convergence observed in the RF algorithm affirm its robustness across diverse datasets.

Nonetheless, the model is not without its limitations. For instance, the RF may falter with excessively large or small datasets; the former may lead to prohibitive computational demands, while the latter could precipitate overfitting. Additionally, the interpretability of RF models may pose challenges due to the inherent complexity of multiple DTs, which obscures the decision-making process compared to more straightforward models. When handling datasets with numerous categorical features, RF may also necessitate specialized preprocessing to maintain performance.

In practical applications of talent cultivation and assessment, the RF model can be employed to significant effect. Educational institutions might utilize it to gauge student engagement and academic performance, thereby forecasting academic success or professional trajectory. Within corporate environments, it could assist HR departments in appraising employee achievements and identifying leaders or talents for further development.

The RF model can further analyze online educational platform data, identifying student challenges and fostering personalized support. It might also serve in the development and assessment of training courses, enabling educators to refine course content in alignment with learning outcomes.

In sum, the RF algorithm stands as a potent, versatile instrument in the appraisal of talent cultivation models, adept at managing intricate datasets and deriving meaningful insights. Despite its strengths, an awareness of the algorithm's limitations is crucial, necessitating tailored adjustments and enhancements to suit the particular demands of varied applications.

Future research could delve into enhancing interpretability, optimizing feature engineering, extending real-time data handling, exploring multitask learning, integrating diverse algorithms, and assessing cross-domain applicability. Pursuing these avenues promises to deepen our understanding of talent cultivation efficacy and sharpen the precision of predictive models in this domain.

Acknowledgments

This work was supported by “Key Education Reform Project at Nanjing University of Technology (20230021)”, “General Project for Philosophy and Social Science Research in Jiangsu Province's Universities (2023SJYB0211)”

Reference

[1] Anlesinya A, Amponsah-Tawiah K. Towards a responsible talent management model[J]. European Journal of Training and Development, 2020, 44(2/3): 279-303.

[2] Zhang J, Gu L. Research on the Optimization of Landscape Design Talent Training Mode Under the Background of Rural Revitalization Strategy [J]. Revista De Gestao Social E Ambiental, 2022, 16(2): e02979-e02979.

[3] Pastrana B, Tobon S. Emerging needs of human talent training in leading information technology companies, a socioformative analysis[J]. World Review of Science, Technology and Sustainable Development, 2020, 16(4): 303-319.

[4] Zhao G, Fan M, Yuan Y, et al. The comparison of teaching efficiency between virtual reality and traditional education in medical education: a systematic review and meta-analysis[J]. Annals of translational medicine, 2021, 9(3).

[5] Zhu W, Lou Y. Research on 3D technology in the field of education: How to make up for the shortcomings of traditional education[J]. Frontiers in Business, Economics and Management, 2022, 6(2): 145-148.

[6] Kibria M G, Nguyen K, Villardi G P, et al. Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks[J]. IEEE access, 2018, 6: 32328-32338.

[7] Rahmani A M, Azhir E, Ali S, et al. Artificial intelligence approaches and mechanisms for big data analytics: a systematic study[J]. PeerJ Computer Science, 2021, 7: e488.

[8] Bharadiya J P. A comparative study of business intelligence and artificial intelligence with big data analytics[J]. American Journal of Artificial Intelligence, 2023, 7(1): 24.

[9] Biau G, Scornet E. A random forest guided tour[J]. Test, f016, f5: 197-ff7.

[10] Sarica A, Cerasa A, Quattrone A. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review[J]. Frontiers in aging neuroscience, f017, 9: 3f9.

[11] Liu Y. Random forest algorithm in big data environment[J]. Computer modelling & new technologies, f014, 18(1fA): 147-151.

[12] [1f] Sun Z, Gao M, Jiang A, et al. Incomplete data processing method based on the measurement of missing rate and abnormal degree: Take the loose particle localization data set as an example[J]. Expert Systems with Applications, f0f3, f16: 119411.

[13] Ruiz-Shulcloper J. Pattern recognition with mixed and incomplete data[J]. Pattern Recognition and Image Analysis, f008, 18: 563-576.

[14] Kanza Y, Nutt W, Sagiv Y. Querying incomplete information in semistructured data[J]. Journal of Computer and System Sciences, f00f, 64(3): 655-693.

[15] Lan Y, Zhang L, Liu L. A method for extracting rules from incomplete information system[J]. Electronic Notes in Theoretical Computer Science, f003, 8f(4): 310-317.

[16] Raviv D. Innovative thinking: Desired skills and related activities[C]//f008 Annual Conference & Exposition. f008: 13.750. 1-13.750. 40.

[17] Drapeau P. Sparking student creativity: Practical ways to promote innovative thinking and problem solving[M]. ASCD, f014.

[18] Ranabahu N, Almeida S, Kyriazis E. University-led internships for innovative thinking: a theoretical framework[J]. Education+ Training, f0f0, 6f(3): f35-f54.

[19] Davis K A, Amelink C T. Exploring differences in perceived innovative thinking skills between first year and upperclassmen engineers[C]//f016 IEEE Frontiers in Education Conference (FIE). IEEE, f016: 1-7.

[20] Kokotsaki D, Menzies V, Wiggins A. Project-based learning: A review of the literature[J]. Improving schools, f016, 19(3): f67-f77.

[21] Krajcik J S, Blumenfeld P C. Project-based learning[M]. na, f006.

[22] Condliffe B. Project-Based Learning: A Literature Review. Working Paper[J]. MDRC, f017.

[23] Tayebinik M, Puteh M. Blended Learning or E-learning?[J]. arXiv preprint arXiv:1306.4085, f013.

[24] Graham C R. Blended learning systems[J]. The handbook of blended learning: Global perspectives, local designs, f006, 1: 3-f1.

[25] Heinze A, Procter C T. Reflections on the use of blended learning[J]. f004.

[26] Gecer A, Dag F. A blended learning experience[J]. Educational Sciences: Theory and Practice, f01f, 1f(1): 438-44f.

[27] Bishop J, Verleger M A. The flipped classroom: A survey of the research[C]//f013 ASEE Annual Conference & Exposition. f013: f3.1f00. 1-f3.1f00. 18.

[28] Akgayir G, Akgayir M. The flipped classroom: A review of its advantages and challenges[J]. Computers & Education, f018, 1f6: 334-345.

[29] Mok H N. Teaching tip: The flipped classroom[J]. Journal of information systems education, 2014, 25(1): 7.

[30] Lurie S J. History and practice of competency-based assessment[J]. Medical education, 2012, 46(1): 49-57.

[31] Harris P, Bhanji F, Topps M, et al. Evolving concepts of assessment in a competency-based world[J]. Medical teacher, 2017, 39(6): 603-608.

[32] Ion G, Cano E, Cabrera N. Competency assessment tool (CAT). The evaluation of an innovative competency-based assessment experience in higher education[J]. Technology, Pedagogy and Education, 2016, 25(5): 631-648.

[33] Malekipirbazari M, Aksakalli V. Risk assessment in social lending via random forests[J]. Expert Systems with Applications, 2015, 42(10): 4621-4631.

[34] Dai B, Chen R C, Zhu S Z, et al. Using random forest algorithm for breast cancer diagnosis[C]//2018 International symposium on computer, consumer and control (IS3C). IEEE, 2018: 449-452.

[35] Gou R, Zhao J. Eco-environmental quality monitoring in Beijing, China, using an RSEI-based approach combined with random forest algorithms[J]. IEEE Access, 2020, 8: 196657-196666.