Talent Quality of Postgraduates in Ship and Ocean Engineering Relying on Improved Random Forest Algorithm

Frey Davide[*]

University of Turin, Italy

[*]*corresponding author*

Abstract: In recent years, the number of graduate students in ship and ocean engineering has been increasing, and the quality of talents as the basic work of talent training determines the starting point of master education in ship and ocean engineering. Due to the gradual expansion of the number of postgraduates, the quality evaluation of postgraduates in ship and ocean engineering has become a top priority. The purpose of this paper is to study the quality of graduate students in ship and ocean engineering relying on improved random forest algorithm. In the experiment, the original data was obtained and the improved random forest algorithm was used. Among the respondents of the questionnaire, the third-year students and postgraduate students who participated in the internship were selected in a targeted manner. By analyzing the quality of talents in the research province, Guarantee the quality of graduate students in ship and ocean engineering.

1. Introduction

The quality assurance system of personnel training for postgraduates in shipbuilding and marine engineering in my country is not yet perfect, and the system and management level cannot fully meet the training requirements of compound talents [1]. It is necessary to objectively understand and rationally analyze the influencing factors such as the quality of students, teachers, facilities, management systems, etc., in order to build a complete set of scientific and complete quality assurance system for the training of postgraduates in ship and ocean engineering.

The quality of students is the basis of education quality, and it is related to whether the established educational goals can be successfully achieved. This is especially true for the master's education in ship and ocean engineering, which aims to cultivate high-level talents. It has clear requirements and standards for students. Papineni S showed in his research that in order to achieve

an efficient organization of a company in terms of human resources, the proposed system designs a model with the help of random forest algorithm by considering different employee parameters. This helps HR to retain employees by identifying gaps and helps the organization run smoothly with good employee retention. The combination of HR and data science can help improve the productivity, collaboration, and well-being of an organization's employees [2]. Azhar Y has researched that customer cancellations of hotel reservations greatly influence hotel management decisions. In order to minimise the damage caused by this issue, hotel management has instituted a rather strict policy that can damage reputation and operating performance. Therefore, research focuses on solving these problems using machine learning algorithms. To obtain the best model performance, hyperparameter optimization is applied to the random forest algorithm. Its purpose is to obtain the optimal combination of model parameters when predicting hotel reservation cancellations. The proposed model proved to have the best performance. The findings can be used as model components related to future booking cancellations in a hotel management decision-making system [3]. If the quality of students is not guaranteed, it will inevitably affect the quality of high-level personnel training.

This paper studies the research status of random forest, clarifies the meaning of student quality and random forest algorithm, and finally establishes a talent quality assurance system for postgraduates in ship and ocean engineering in universities, from strict control of student source quality input assurance and a serious view of external evaluation output to ensure the quality of talents system. In the experiment, the original data was obtained and the improved random forest algorithm was used. Among the respondents of the questionnaire, the third-year students and postgraduate students who participated in the internship were selected in a targeted manner. By analyzing the quality of talents in the research province, Guarantee the quality of graduate students in ship and ocean engineering.

2. Research on Talent Quality of Postgraduates in Ship and Ocean Engineering Relying on Improved Random Forest Algorithm

2.1. Quality of Students

There is no authoritative discussion about the meaning of "quality of students". In real life, people's understanding of quality of students is a process from shallow to deep [4-5]. At first, it was believed that the quality of students was equal to the scores of students when they entered the school, and the score represented the quality of students. Later, some scholars proposed that the quality of students should include not only the grades of students when they enter the school, but also the basic knowledge of the students. At this stage, the connotation of student quality has been continuously improved, which has increased the performance of candidates in terms of admission scores, professional knowledge base, learning attitude, and psychological quality [6-7]. Quality expresses the degree to which stated and implicit needs are met. According to this understanding, the understanding of the meaning of the quality of students should focus on the extent to which the candidates can meet the improvement needs of the school. Under the current situation that higher education is becoming more and more popular, it is also an important content to grasp the connotation of the quality of students whether to meet the needs of the society as much as possible. At present, there is a great controversy about the "quality of students' source" in the academic circles. From a macro perspective, some researchers understand the quality of students as the degree to which a certain level of knowledge and skills acquired by students at the beginning of a certain educational stage meet the requirements of education and training at that stage [8-9].

2.2. Random Forest Algorithm

The random forest algorithm is a classification algorithm that integrates multiple decision trees [10-11]. Random forest has several advantages in machine learning: it adopts ensemble algorithm, which has high accuracy; it can process high-dimensional data, etc., and has strong adaptability to data sets; the training speed is fast, and it can be used in large-scale data sets [12- 13].

The decision tree is the base classification tree of the random forest algorithm. Two randomization processes are introduced on the basis of the base classification tree: the process of randomly extracting samples with replacement; the process of randomly extracting features without replacement. The mainstream algorithms of decision tree include ID3 algorithm, C4.5 algorithm and CART algorithm [14-15]. Decision tree, also known as decision tree, is a basic instance-based classification and regression method. It is mainly composed of nodes and directed edges. The nodes include root nodes, internal nodes and leaf nodes. Generally, a decision tree contains root nodes, internal nodes, and leaf nodes, which respectively represent the complete set of samples, an attribute or feature, and the decision result (a data set that has been classified). According to the test result of the attribute, the included sample set is divided into child nodes. A directed edge is a path, representing a classification rule, and each path is mutually exclusive and complete. The top ellipse represents the root node, the remaining ellipses represent internal nodes, and the rectangles are leaf nodes. A decision tree is shown in the simple decision tree legend in Figure 1:
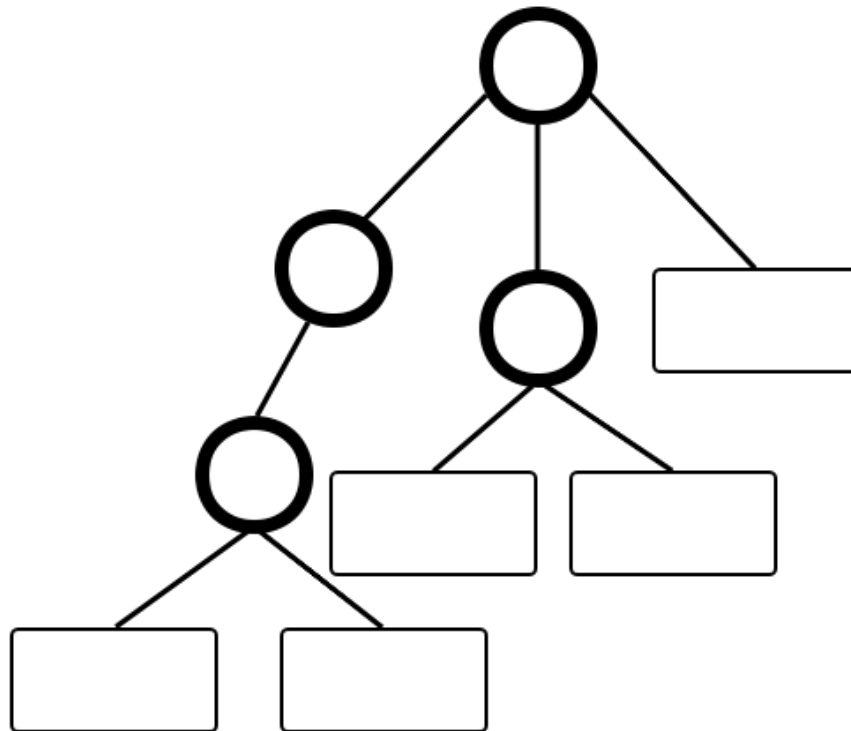

Figure 1. Simple decision tree legend

2.3. Quality Assurance System for Graduate Students in Shipbuilding and Marine Engineering in Colleges

(1) Strictly control the quality input of students

The quality of student sources, as a component of "input" in the process of cultivating graduate students studying abroad, has also become the top priority in the three dimensions of "input, process, and output"[16-17]. As far as the quality of the national graduate students in ship and ocean engineering is concerned, the quality of the students is in a state of mixed quality. Therefore, major universities cannot continue to blindly expand enrollment. The economic dividends of students and the rigid achievement of government indicators have caused the reputation of Chinese universities to fail, so keep the appropriate scale for its number. Secondly, when selecting the best students for the quality of their students, they should be cautious about the quality of their thoughts.

(2) Take the external evaluation output guarantee seriously

As a component of "output" in the process of cultivating graduate students studying abroad, external evaluation occupies a very high position in the three dimensions of "input, process, and output". This process also directly sends it to the society and accepts it. evaluation [18]. The comprehensive weight of external evaluation is the largest, so in the first-level indicators, external evaluation plays an important role in it. In order to improve the quality assurance system for the training of postgraduates in ship and ocean engineering in universities. External evaluations of graduate students in ship and marine engineering should be taken seriously. The job remuneration, job competency, performance and improvement potential in the social evaluation can be considered as a direct reflection of the quality of training, which will be further fed back to the quality of students, training conditions, and training results, so that the first-level indicators can be further obtained. Improve and get positive social evaluation.

3. Investigation and Research on Talent Quality of Postgraduates in Ship and Ocean Engineering Relying on Improved Random Forest Algorithm

3.1. Acquisition of Raw Data

Among the respondents of the questionnaire, two types (in school and working) and three levels (better, middle, and poor) of college students were selected in a targeted manner: the third-year graduate students who participated in the internship D , Graduate student Z.

3.2. Improved Random Forest Algorithm

Based on the random forest of the C4.5 algorithm, in the testing phase, each value that needs to be tested goes through all the trees in the forest at the same time, starting from the root node until it reaches the corresponding leaf node. The predicted value of the RF algorithm is based on the majority vote of each tree, i.e. the class with the most votes is predicted by the tree. In order to obtain the output prediction of the RF algorithm, according to the i-th observation value of the test data set, the output prediction value of the tree $p(p = 1,...,p)$ is expressed as y_{ip}, y_i represents the prediction output of the entire RF algorithm, which can be calculated by the formula:

$$y_i = \frac{1}{p}\sum_{i=1}^{p} y_{ip}$$

(1)

The classification results based on y_i 's hypothesis testing are implemented as follows:

$$\left\{ \begin{array}{ll} LOS & if\ y_i \geq 0 \\ NLOS & if\ y_i\ \pi\ 0 \end{array} \right\}$$

(2)

4. Analysis and Research on Talent Quality of Postgraduates in Ship and Ocean Engineering Relying on Improved Random Forest Algorithm

Taking the third-year graduate student D and the current graduate student Z who participated in the internship as the research objects, the selected candidates obtained the opinions of each stakeholder by asking their parents, teachers, student work managers, employer leaders, and colleagues for their views. The evaluation of various indicators on themselves, that is, the utility value, and the weighted average value is calculated for each type of utility value. Among them, indicators C1 (basic knowledge of the major), C2 (cultural accomplishment (such as natural knowledge, humanistic knowledge, etc.)), C3 (law, legal awareness and ideological and moral quality), C4 (honest and trustworthy, hard-working, pragmatic and hard-working) style), C5 (physical fitness). The data summary is shown in Table 1 and Figure 2, Table 2 and Figure 3:

Table 1. Each utilit's weighted average of student D

Various indicators	C1	C2	C3	C4	C5
Self	0.9	0.5	0.8	0.9	0.7
The head of a family	0.8	1	0.7	0.8	0.7
Teacher	1	1	0.7	0.8	0.7
A Student Work Manager	0.7	0.8	0.9	0.6	1
Employer	0.8	0.7	0.6	0.5	0.9

Figure 2. Evaluation diagram of the third students participating in the internship

Table 2. The weighted mean of each utility for the Z of the school graduate students

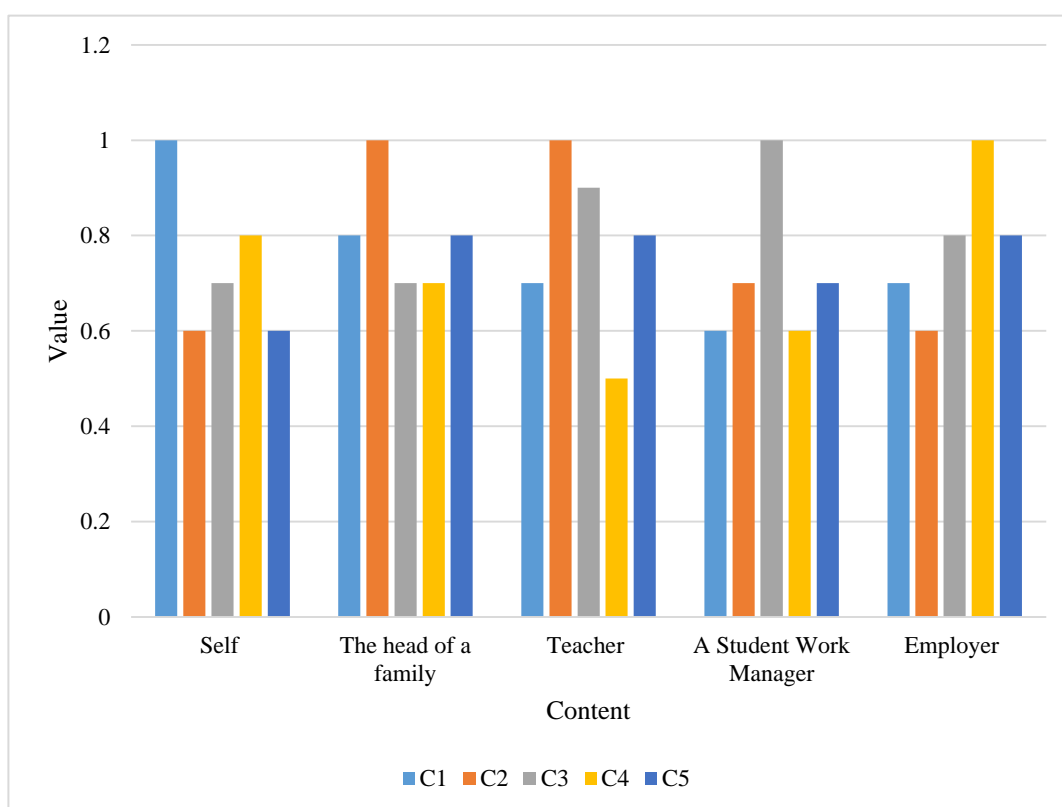| Various indicators	C1	C2	C3	C4	C5
Self	1	0.6	0.7	0.8	0.6
The head of a family	0.8	1	0.7	0.7	0.8
Teacher	0.7	1	0.9	0.5	0.8
A Student Work Manager	0.6	0.7	1	0.6	0.7
Employer	0.7	0.6	0.8	1	0.8

Figure 3. Z-evaluation diagram of the school graduate students

After careful comparative analysis, it can be found that the lowest utility value is Z is 0.574; the highest utility value is D is 0.985. According to this, give guidance for their future efforts and stage goals respectively, Z needs to pay special attention to strengthening the learning and application of professional knowledge, cultivate innovation ability, and pay attention to the practice of ideological and moral quality and good work style; D needs solid lifelong learning The concept of continuous learning, combined with practice, to make yourself better. In order to continuously improve the quality of graduate students in ship and ocean engineering.

5. Conclusion

In order to improve the quality assurance system for the training of postgraduates in ship and ocean engineering in universities, quality is not only the lifeline of domestic postgraduate training, but also the lifeline of postgraduate training in China. Although with the in-depth reform of postgraduate education in my country, the government's emphasis on the quality assurance of

domestic postgraduate training has gradually increased, but the emphasis on the quality assurance of postgraduate training in ship and ocean engineering is somewhat insufficient. It should be based on the goal of the quality of postgraduates in ship and ocean engineering, that is, the goal of cultivating postgraduate talents in ship and ocean engineering, combined with the current relevant research, to further improve the quality assurance of the training of postgraduates in ship and ocean engineering in universities system.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Urbanovich E A , Afonnikov D A , Nikolaev S V . Determination of the quantitative content of chlorophylls in leaves by reflection spectra using the random forest algorithm. Vavilov Journal of Genetics and Breeding, 2021, 25(1):64-70. https://doi.org/10.18699/VJ21.008

[2] Papineni S , Reddy A M , Yarlagadda S , et al. An Extensive Analytical Approach on Human Resources using Random Forest Algorithm. International Journal of Engineering Trends and Technology, 2021, 69(5):119-127. https://doi.org/10.14445/22315381/IJETT-V69I5P217

[3] Azhar Y , Mahesa G A , Mustaqim M C . Prediction of hotel bookings cancellation using hyperparameter optimization on Random Forest algorithm. Jurnal Teknologi dan Sistem Komputer, 2021, 9(1):15-21. https://doi.org/10.14710/jtsiskom.2020.13790

[4] Onesime M , Yang Z , Dai Q . Genomic Island Prediction via Chi-Square Test and Random Forest Algorithm. Computational and Mathematical Methods in Medicine, 2021, 2021(1):1-9. https://doi.org/10.1155/2021/9969751

[5] Lacerda E R , Vicens R S . Detection Of Invariant Vegetation Areas In Time Series Using Random Forest Algorithm. GEOgraphia, 2021, 23(50):1-12. https://doi.org/10.22409/GEOgraphia2021.v23i50.a46996

[6] Pang C . Simulation of student classroom behavior recognition based on cluster analysis and random forest algorithm. Journal of Intelligent and Fuzzy Systems, 2021, 40(2):2421-2431. https://doi.org/10.3233/JIFS-189237

[7] Kim J Y , Lee M , Min K L , et al. Development of Random Forest Algorithm Based Prediction Model of Alzheimer's Disease Using Neurodegeneration Pattern. Psychiatry Investigation, 2021, 18(1):69-79. https://doi.org/10.30773/pi.2020.0304

[8] Ramalingam S , Baskaran K . An efficient data prediction model using hybrid Harris Hawk Optimization with random forest algorithm in wireless sensor network. Journal of Intelligent and Fuzzy Systems, 2020, 40(20):1-25. https://doi.org/10.3233/JIFS-201921

[9] Klnarslan E , YI Türker, Nce M . Prediction of Heat-Treated Cedar Wood Swelling And Shrinkage Values With Artificial Neural Networks and Random Forest Algorithm. Mühendislik Bilimleri ve Tasarım Dergisi, 2020, 8(5):200-205. https://doi.org/10.21923/jesd.825442

[10] Vimala K , Usha D . An Efficient Classification of Congenital Fetal Heart Disorder using Improved Random Forest Algorithm. International Journal of Engineering Trends and Technology, 2020, 68(12):182-186. https://doi.org/10.14445/22315381/IJETT-V68I12P229

[11] Pasinetti S , Fornaser A , Lancini M , et al. Assisted Gait Phase Estimation Through an Embedded Depth Camera Using Modified Random Forest Algorithm Classification. IEEE Sensors Journal, 2020, 20(6):3343-3355. https://doi.org/10.1109/JSEN.2019.2957667

[12] Rewade A D , Mohod S W , Bargat S P . Content Based Alternate Medicine Recommendation By Using Random Forest Algorithm. International Journal Of Computer Sciences And Engineering, 2019, 7(4):1163-1168. https://doi.org/10.26438/ijcse/v7i4.11631168

[13] Han S S , Kim M S , Lim W , et al. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. Journal of Investigative Dermatology, 2018, 138(7):1529-1538. https://doi.org/10.1016/j.jid.2018.01.028

[14] Grassmann F , Mengelkamp J , Brandl C , et al. A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography. Ophthalmology, 2018, 125(9):1410-1420. https://doi.org/10.1016/j.ophtha.2018.02.037

[15] Bourassa A E , Roth C Z , Zawada D J , et al. Drift corrected Odin-OSIRIS ozone product: algorithm and updated stratospheric ozone trends. Atmospheric Measurement Techniques, 2018, 11(1):489-498. https://doi.org/10.5194/amt-11-489-2018

[16] Aljarah I , Faris H , Mirjalili S . Optimizing connection weights in neural networks using the whale optimization algorithm. Soft Computing, 2018, 22(1):1-15. https://doi.org/10.1007/s00500-016-2442-1

[17] Gong D W , Jing S , Miao Z . A Set-Based Genetic Algorithm for Interval Many-Objective Optimization Problems. IEEE Transactions on Evolutionary Computation, 2018, 22(99):47-60. https://doi.org/10.1109/TEVC.2016.2634625

[18] Primitivo B. Acosta-Humánez a, J. Tomás Lázaro b, C M R , et al. Differential Galois theory and non-integrability of planar polynomial vector fields. Journal of Differential Equations, 2018, 264(12):7183-7212. https://doi.org/10.1016/j.jde.2018.02.016