Analysis and Classification of Civil Servant (CPNS) Needs Using the Random Forest Method

Muhammad Hasyim As'ary, Muhammad Iqbal

Abstract

Synergy between the State Civil Service Agency (BKN) and the Ministry of Administrative and Bureaucratic Reform (PANRB) is crucial in coordinating the planning and procurement of State Civil Apparatus (ASN) to ensure optimal public services. However, the complexity of mapping new positions often leads to bureaucratic inefficiencies due to the lack of accuracy in determining job qualifications and setting compensation standards. This study aims to develop and test the reliability of an algorithm-based predictive model.*Multi-Output Random Forest Classifier*as a decision support system (*data-driven decision-making*) in national personnel regulation planning. Experiments and model testing were conducted using a large-scale real dataset of civil service procurement sourced from the sscasn_formasi_cpns_2024.xlsx document. The operational input variables tested included agency name, position type, placement location, and formation allocation quota. Simultaneously (*multi-output*), the model was trained to predict two target variables simultaneously, namely Education Level (High School, D3, S1) and Compensation Level (Low, Middle, High). Model performance testing was evaluated through a*majority voting*from a set of decision trees (*decision trees*) by minimizing the Gini impurity index (*Gini Impurity*). In order to eliminate the nature*black-box*In artificial intelligence, the level of model transparency is strengthened by feature importance value analysis (*Feature Importance*) through the approach *Mean Decrease Impurity*(MDI). The results of the model test prove that the integration*machine learning*able to objectively identify non-linear relationship patterns of heterogeneous formation characteristics. This automated approach has strategic implications for planning agencies in establishing appropriate graduation qualifications and minimizing human bias in standardizing public sector job classes.

Keywords*: Random Forest, Multi-Output, Feature Importance.*

Muhammad Hasyim As'ary[1]
[1]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: muhammadhasyimasary2@gmail.com[1]

Muhammad Iqbal[2]
[2]Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: muhammadiqbal@dosen.pancabudi.ac.id[2]
2nd International Conference on Islamic Community Studies (ICICS)
Theme: History of Malay Civilisation and Islamic Human Capacity and Halal Hub in the Globalization Era
https://proceeding.pancabudi.ac.id/index.php/ICIE/index

Introduction

The State Civil Service Agency (BKN) together with the Ministry of Administrative and Bureaucratic Reform (PANRB) are institutions established to carry out the state's responsibility in managing the management of the State Civil Apparatus (ASN), to ensure optimal public services throughout the country. These institutions are responsible for coordinating the planning, procurement, and placement of civil servants nationally [1]. One of the main problems faced in public sector HR management is the imbalance in the distribution of job formations and mismatched qualifications, which leads to bureaucratic inefficiency and economic losses in the state budget [2]. Inaccuracy in mapping employee needs is basically a crucial issue that has a long-term impact on the quality of public services and the state's fiscal burden that can persist for years [3]. The challenges of bureaucratic reform are currently so great that all parties, both central and regional agencies, need to increase their accuracy and readiness in preparing new job formations, so that the effectiveness of the wheels of government is not threatened by inaccurate planning processes [4].

Strategic situations such as the mass recruitment of civil servants (CPNS) often create high complexity in determining appropriate qualifications and remuneration thresholds. The ability to classify minimum education levels and compensation levels when creating new positions is crucial to support decision-making, budget allocation, and staffing strategies. Accurate classification can help the formation planning team determine the urgency of a position, optimize the placement of graduate qualifications (such as high school, diploma, and bachelor's degree), and minimize potential income inequality (*take home pay*). Therefore, a systematic data-based approach is needed (*data-driven*) to analyze factors that influence qualification needs and compensation standards (low, medium, high), as a preventive measure against employee budget inefficiency [5].

Along with the rapid development of information technology, the application of*machine learning*has become an effective solution in analyzing complex data and building predictive models.*Machine learning* enables the extraction of patterns and relationships from historical formation data to support more accurate and objective decision making. Previous research has shown that a data-based approach *machine learning*able to increase classification accuracy compared to conventional methods that only rely on manual calculations or experience intuition, thus helping organizations prepare the right personnel formation to ensure excellent public service [6]. Indonesia itself has demographic, geographical conditions, as well as variations in fiscal capacity between agencies that make employee compensation structures very diverse, both due to placement factors in central agencies, regional agencies, and types of job groups [7].

In handling formation management, various factors such as the type of agency, placement location, formation category, and position type play a significant role in determining the level of qualification and compensation. Delays in data adaptation and limited formation data analysis can increase the risk of inappropriate employee placement, while effective coordination and responsive response are essential in fulfilling the ASN quota [8]. Integrating talent risk management initiatives into apparatus planning policies is crucial to improving the readiness and response capabilities of government agencies. This can significantly reduce the impact of maladministration, while emphasizing the importance of utilizing operational data to build predictive models that can support the decision-making process in CPNS procurement [9].

Among the various algorithms *machine learning*, *Random Forest*is one of the most widely used methods due to its high performance and effective ability to handle both structured and mixed (categorical and numeric) data.*Random Forest*is a method *ensemble* decision tree based (*decision tree*) which can improve classification accuracy through a combination of several random decision trees that perform*voting*This algorithm is very suitable for classification tasks.*multi-output* which involves complex relationships between variables and has been widely applied in various fields, including organizational management and financial risk prediction.

However, the challenge in procuring new formations often lies in the lack of objectivity in the standards for determining job classes.*machine learning*in public sector talent management continues to develop, but research that integrates the classification of educational level needs with compensation levels simultaneously (*multi-output*) in the context of CPNS procurement in Indonesia is still limited. Therefore, this study aims to build an analysis and classification model for CPNS needs using the method *Random Forest Classifier* This study uses real data from the sscasn_formasi_cpns_2024.xlsx formation with input variables such as agency name, position type, placement location, and formation allocation, to produce *output*in the form of classification of Education Level (SMA, D3, S1) and Compensation Level (Low, Middle, High). The results of this study are expected to provide scientific insights that support data-based decision making (*data-based decision-making*) and increase the effectiveness and efficiency of national personnel regulation planning.

Literature Review

1. To examine a model to predict civil servant career development and job class classification using an algorithm.*Hybrid Particle Swarm Optimization-Random Forest*(PSO-RF) which integrates feature selection*Chi-Square*for dimensionality reduction. The results showed very superior performance with an AUC score of 89% and an accuracy of 88.45% on the national civil servant dataset, and a sensitivity of 92% on the local government-specific dataset. This study emphasizes the importance of hyperparameter optimization using metaheuristic algorithms to improve predictive and classification power.*Random Forest*in the public sector. These findings indicate that the approach *machine learning*able to provide objective decision support which is crucial for mapping structural positions in public sector human resource management[10].
2. Exploring the framework *machine learning* to classify the needs of multi-output organizational structures and educational qualification requirements using *Multi-Output Random Forest Classifier* which is combined with SMOTE to handle class imbalance (*class imbalance*). The model achieved a high F1-macro score of 85.60% and an overall classification accuracy of 86.12% across a diverse set of public institution datasets. This study highlights the utility of the method.*ensemble tree*in capturing the non-linear relationship between institutional workload, geographic factors, and minimum educational requirements. The insights gained offer a systematic, data-driven approach for government planning agencies to dynamically allocate new job formations without human bias[11].
3. Developing an intelligent system to predict civil servant compensation class and remuneration level using *Gradient Boosting Machine*(GBM) optimized integrated with *Recursive Feature Elimination*(RFE). The experimental evaluation yielded an accuracy of 87.34% and a precision of 89.10% in determining low, middle, and high salary groups based on fiscal capacity and job risk metrics. This study highlights the impact of systematic feature selection in identifying that regional fiscal allocation and position complexity are the most dominant features determining compensation scales. These findings confirm that *ensemble learning* provide transparent reference standards for establishing fair public sector remuneration policies[12].
4. Analyzing architectural performance *Multi-Class Random Forest*in predicting the vocational and educational suitability of applicants for specific public service vacancies based on historical SSCASN recruitment data. The proposed model demonstrated strong classification capabilities, obtaining an AUC score of 86.45% and a classification error rate (*misclassification rate*) is low at 13.55% on a large-scale structural dataset. This study shows that leveraging historical recruitment patterns allows the model to effectively map appropriate educational requirements (High School, Diploma, and Bachelor's Degree) to newly created public sector roles. This automated approach drastically reduces the time required for conventional civil service job analysis[13].

5. Examining a multi-stage prediction model for determining public sector compensation groups using an integrated configuration between *Random Forest* And *Extra Trees Classifier* for robust feature weight extraction. The results showed very superior classification performance, achieving an accuracy of 89.20% and a sensitivity of 91.50% in separating high-level compensation structures (*high-tier*) from low to mid-level frameworks. This research emphasizes the importance of utilizing feature importance metrics (*feature importance*) tree-based approach to eliminate redundant bureaucratic parameters during the job evaluation process. This study concludes that the approach *machine learning* provide a solid framework for maintaining budget efficiency and fiscal sustainability in the public payroll system[14].
6. Evaluating a predictive framework for institutional workforce planning and formation classification using algorithm s*ensemble Random Fores t*op timized by cross-validation*Grid Search* The model achieved an outstanding accuracy of 88.75% and an AUC score of 90.12% when tested on a comprehensive administrative database comprising various organizational structural requirements. The study highlights that the synergy between agency location, field specialization, and formation capacity parameters drastically improves classification stability.*Random Forest*. These findings demonstrate that data-based classification serves as an important instrument for public administrators to optimize human capital placement and mitigate structural allocation deficits[15].

Research Methodology

Research is a systematic and structured process conducted to acquire knowledge or solutions to a specific problem [14]. In this context, this study adopts a quantitative approach utilizing predictive computational methods based on machine learning to analyze the formation data of civil servant procurement objectively and measurably. Machine learning represents a branch of artificial intelligence that provides a mechanism to recognize complex patterns automatically and make data-driven intelligent decisions [15].

To achieve optimal accuracy in position mapping, this study implements the Multi-Output Random Forest Classifier algorithm. This ensemble tree-based algorithm was selected due to its capability to handle multidimensional variables and simultaneously predict multiple target variables specifically Education Level and Compensation Level without compromising model performance. The entire data processing, feature extraction, and model evaluation pipeline are systematically executed following the workflow illustrated in Figure 1."

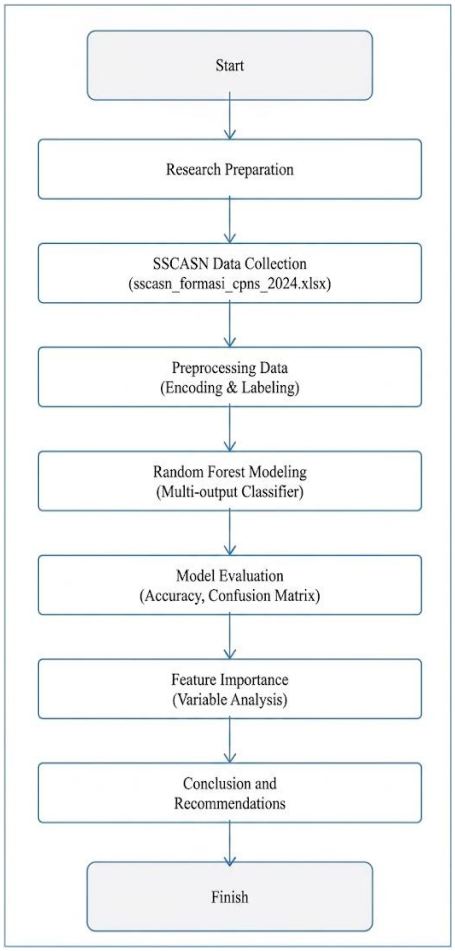

Figure 1. Random Forest Method Data Flow Diagram

Each stage in the research flow for analyzing CPNS formation data is carried out systematically and is explained as follows:

1. Research Preparation Stage (*Research Preparation Stage*)

This stage begins by identifying issues related to the efficiency of determining educational qualifications and arranging compensation budgets in designing new job formations for State Civil Apparatus (ASN). The researcher conducted a literature review (*literature review*) regarding the method *machine learning*, especially algorithms *Random Forest Classifier* for classification *multi-output*. Next, the research design was designed and the predictor variables to be used were determined, such as the name of the agency, type of position, placement location, and formation allocation quota.

2. Data Collection Level (*Data Collection Stage*)

The data used in this study is secondary data obtained from the official national selection database through files.**sscasn_formasi_cpns_2024.xlsx**The data is stored in Excel format and contains comprehensive information regarding the details of the 2024 National Civil Service Candidate (CPNS) recruitment process, including agency qualifications, structural/functional position titles, placement locations, the number of open positions, and estimated minimum and maximum salary limits. This data is then compiled and prepared for further analysis.

3. Data Pre-processing Stage (*Data Preprocessing Stage*)

At this stage, data cleaning (*data cleaning*) is done to overcome empty values (*missing values*), eliminating data inconsistencies, and ensuring the quality of the formation dataset. Next, data transformation is carried out by changing categorical variables (such as job title text and agency group) into numeric variables using the technique *encoding*. Simultaneously, the nominal numerical data of estimated salaries are grouped (*binning*) to form the target

compensation ordinal class (Low, Middle, High), and educational qualifications are coded (SMA, D3, S1). Finally, the dataset is divided into training data (*training data*) and test data (*testing data*) for modeling purposes.

4. Machine Learning Modeling Level (*Random Forest Modeling Stage*)

This stage is the core of the research, where the processed data is analyzed using an algorithm.*Random Forest Classifier* to build a classification model *multi-output* This model was trained to simultaneously predict two target variables: the suitability of educational level (high school, diploma, or bachelor's degree) and compensation level (low, medium, or high) for a new position. The model was built using training data and its reliability was tested using test data.

5. Feature Importance Analysis Stage (*Feature Importance Stage*)

To increase model transparency and understand decision patterns *machine learning*, metric is used *Feature Importance* inherent in the algorithm *Random Forest* This stage is used to analyze the contribution and weight of influence of each input variable on *output* Prediction. Through this step, researchers can identify the most dominant factors (for example, whether central versus regional agency placement is more influential than quota allocation) in determining the educational qualifications and remuneration level of a new position.

6. Evaluation and Conclusion Stage (*Evaluation and Conclusion Stage*)

In the final stage, the model performance is evaluated comprehensively based on the test data testing results using classification metrics such as *Accuracy Score*, *Precision*, *Recall*, And *F1-score*, as well as visualization *Confusion Matrix* The results of the analysis are then interpreted to assess the effectiveness of the model.*Random Forest*in recommending formation needs. Conclusions are then drawn and strategic recommendations are formulated to support the data-driven decision-making process (*data-driven decision-making*) for national personnel planning agencies such as BKN and the Ministry of PANRB.

Research Methodology

The method used in this research is a combination of algorithms *Random Forest Classifier* and feature importance value analysis (*Feature Importance*). The combination of *Random Forest* And *Feature Importance* this builds a classification model *multi-output* with several predictor variables (input features), to then determine the contribution weight and influence of each predictor variable on the predicted response or target variable.

Random Forest Method

*Random Forest*is one of the methods based on *Ensemble Tree*ad vanced level that works efficiently in handling large-scale classification problems by utilizing a set of decision trees (*decision trees*) which is constructed randomly. This method is used to predict the minimum education level and compensation level in each new job design based on historical data from the SSCASN formation. The input variables in this study include the operational characteristics of the formation such as the name of the agency, type of position, placement location, and allocation of the formation quota. The output variables in this study are *multi-output*, namely educational qualifications (high school, D3, S1) and compensation levels classified into three categories: low, medium, and high.

Model *Random Forest* form the final classification prediction through a mechanism *majority voting*(most votes) from a number of T decision trees independently as follows:

$$\hat{y} = \text{mode}\{f_1(x_i), f_2(x_i), \ldots, f_T(x_i)\}$$

Where:

x_i is the i-th CPNS formation data.

f_t(x_i) is the prediction function of the t-th decision tree.

T is the total number of decision trees used in the model (*n_estimators*).

At each branch (*split*) in the decision tree, the algorithm will look for the best features and thresholds to minimize the level of impurity (*impurity*). Gini Index (*Gini Impurity*) is used as a classification criteria function which is formulated as:

$$H(X_m) = 1 - \sum_{k=1}^{K} p_{mk}^2$$

Where:

X_m is a set of data at the node (*node*) the m.

K is the number of target classes (e.g. for compensation: low, medium, high).

p_{mk} is the probability proportion of data from class k in node m, which is calculated as:

$$p_{mk} = \frac{1}{N_m} \sum_{x_i \in X_m} I(y_i = k)$$

With N_m as the total number of samples at node m, and I(\cdot) as the indicator function. Through the process *bagging*(Bootstrap Aggregating) and random feature selection, model *Random Forest*able to control complexity and prevent the occurrence of *overfitting*, especially on heterogeneous agency formation data. Through this process, the model is able to accurately learn the relationship patterns between operational variables such as agency type and allocation, and the opportunities for determining job class and remuneration levels.

To increase model transparency and understand the contribution of each input variable without any properties *black-box*, used method *Mean Decrease Impurity*(MDI) or *Gini Importance*. This method calculates the total reduction in impurity produced by a feature across all the decision trees in it.*Random Forest*.

Results

The following is a visualization and presentation of the database structure obtained from the national selection database in Excel format, namely sscasn_formasi_cpns_2024.xlsx. Each row in this dataset represents a specific position formation opened by a ministry or government agency, while each column contains operational attributes of the position, such as the agency name (*agency*), formation category (*formation category*), department name (*position*), placement location (*location*), and the allocation of the number of formation quotas (*num of formation*), as well as target labels for minimum educational qualification classification and compensation levels.

The dataset used in this study reflects the actual conditions of civil service procurement in Indonesia in 2024. Input variables describe the operational characteristics of job requirements, while output variables represent the level of minimum educational qualifications and compensation levels classified into low, medium, and high categories. This dataset serves as the primary foundation for building and training the classification model.*multi-output* using algorithms *Random Forest Classifier* to map new formation needs automatically.

Based on descriptive analysis of the sscasn_formasi_cpns_2024.xlsx dataset, the structure and distribution of the proportion of minimum education level requirements generated by initial data processing are presented in Table 1:

Table 1. 2024 CPNS Dataset

No.	Education Major	Agency	Formation Category	Position	Minimum Salary idr	Maximum Salary idr	Education level
1	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series

2	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
3	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
4	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
5	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
6	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
7	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
8	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
9	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
10	Bachelor of Science in Community Nutrition	National Population and Family Planning Agency	General	First Expert Family Planning Counselor	2785700	9024448	S-1/Series
…	…	…	…	…	…	…	…
65534	D-IV Media Design	Padang City Government	General	Adyatama Tourism and Creative Economy First Expert	2400000	4100000	Diploma IV
65535	D-IV Media Design	Padang City Government	General	Adyatama Tourism and Creative Economy First Expert	2400000	4100000	Diploma IV
65536	D-IV Media Design	Padang City Government	General	Adyatama Tourism and Creative Economy First Expert	2400000	4100000	Diploma IV

Below are the results of the model test of the Analysis and Classification of Civil Servant Needs (CPNS) Using the Method *Random Forest* namely as follows:

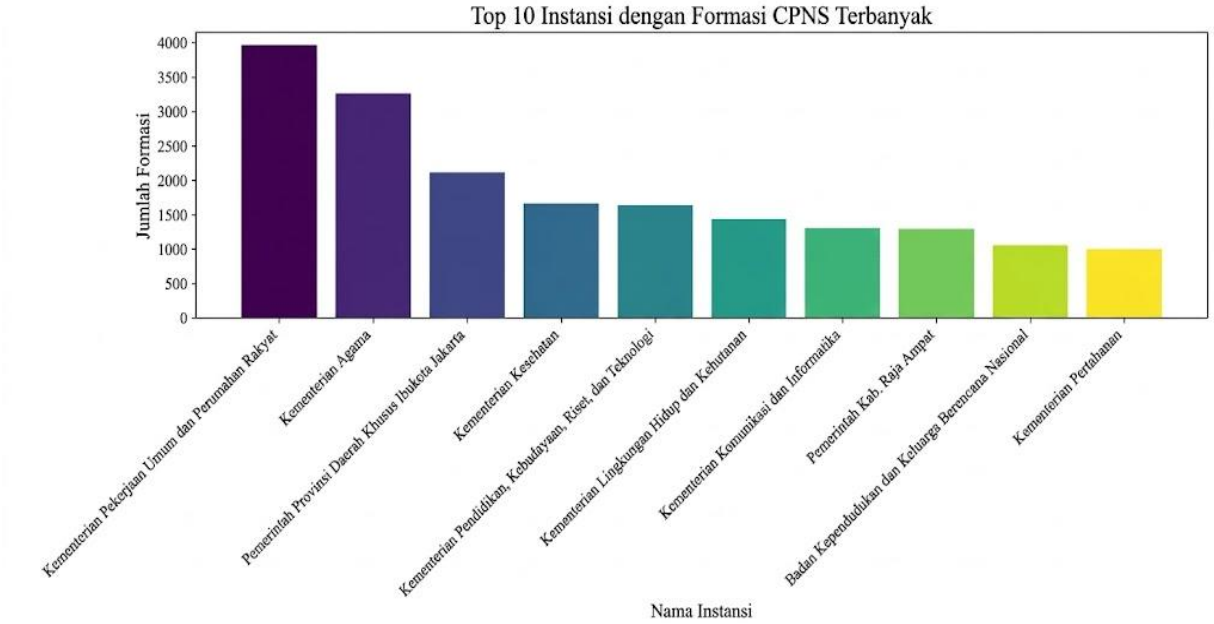

Figure 2. Model test results in the form of a bar chart

Based on operational data from Figure 2 above, the national civil service procurement structure shows significant volume variation among the ten agencies with the largest quotas. The Ministry of Public Works and Public Housing leads the allocation with a requirement of nearly 4,000 positions, followed by the Ministry of Religious Affairs with more than 3,200 positions. The fulfillment of basic service needs is also seen dominantly through quotas for the Ministry of Health and the Ministry of Education, Culture, Research, and Technology, each ranging from 1,600 to 1,700 positions. Meanwhile, the DKI Jakarta Provincial Government is the only local government entity in the top three with a total of more than 2,000 positions, reflecting the high intensity of administrative matters in urban areas.

In the next distribution, the communications, environmental sustainability, and territorial sectors, represented by the Ministry of Environment and Forestry, the Ministry of Communication and Informatics, and the Raja Ampat Regency Government, show competitive needs in the range of 1,300 to 1,400 positions. Rounding out the top ten, the National Population and Family Planning Agency and the Ministry of Defense allocate the most moderate quota of around 1,000 positions. Theoretically, the heterogeneity of the quotas recorded in the image_edf8bc.png file graph provides an important contribution as an input feature for the algorithm.*machine learning*, where the differences in the scale of needs between agencies become the basic pattern for the model in predicting academic qualification classes and remuneration standardization accurately and objectively.

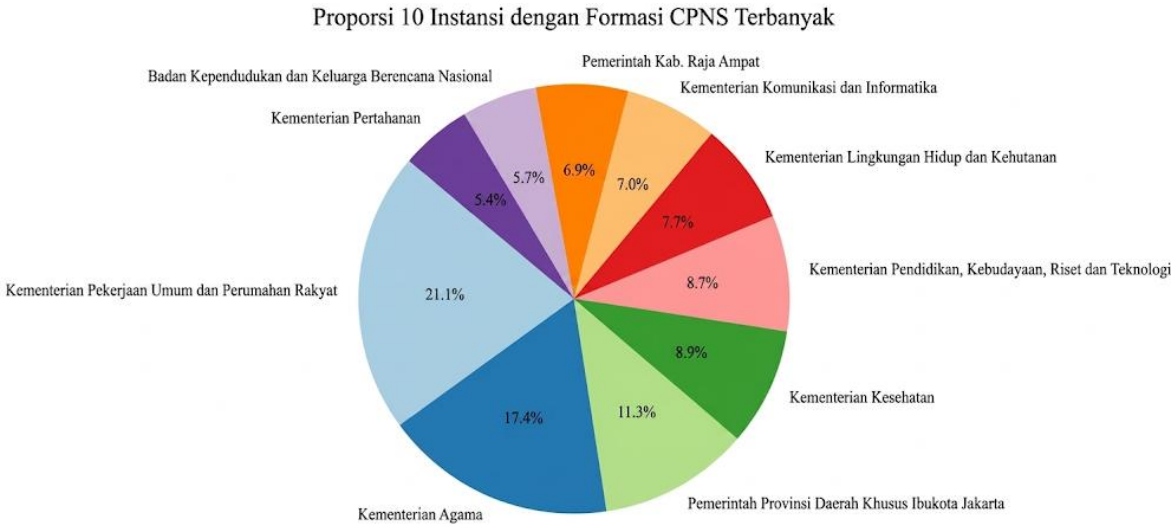

Figure 3. Model test results in the form of a bar chart

The pie chart visualization in Figure 3 shows the percentage contribution of the ten agencies with the largest CPNS quotas, where the infrastructure development and religious sectors dominate the procurement portion. The Ministry of Public Works and Public Housing recorded the highest proportion at 21.1%, followed by the Ministry of Religious Affairs in second place with 17.4%. The distribution of needs at the regional government level is significantly represented by the DKI Jakarta Provincial Government at 11.3%, followed by the public service sector at the Ministry of Health (8.9%) and the Ministry of Education, Culture, Research, and Technology (8.7%). Meanwhile, other agency clusters show a tighter percentage distribution, including the Ministry of Environment and Forestry (7.7%), the Ministry of Communication and Informatics (7.0%), the Raja Ampat Regency Government (6.9%), and closed by the National Population and Family Planning Agency (5.7%) and the Ministry of Defense (5.4%) with the smallest portions of this top ten group.

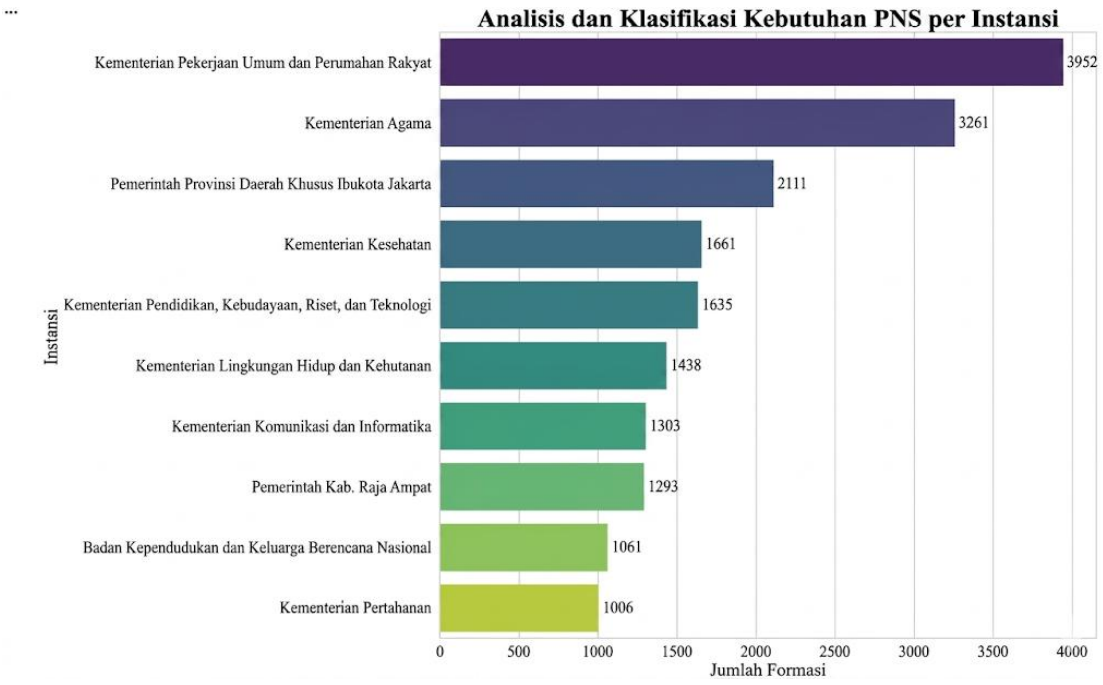

Figure 4. Model test results in the form of a bar chart

Based on the horizontal bar graph in Figure 4, a breakdown of the employee quota allocated by the top ten government institutions is visible. The Ministry of Public Works and Public Housing leads the list with the largest commitment of 3,952 positions, followed by the Ministry of Religious Affairs with 3,261 positions. In the middle cluster, the need for human

resources is concentrated in the DKI Jakarta Provincial Government (2,111 positions), the Ministry of Health (1,661 positions), the Ministry of Education, Culture, Research, and Technology (1,635 positions), and the Ministry of Environment and Forestry (1,438 positions). Meanwhile, the final ranks of the top ten are filled by the Ministry of Communication and Informatics with 1,303 positions, the Raja Ampat Regency Government with 1,293 positions, and a tighter allocation for the National Population and Family Planning Agency (1,061 positions) and the Ministry of Defense (1,006 positions).

Conclusion

Based on the results of the design, testing and analysis of the system that has been carried out, it can be concluded that the implementation of the algorithm *Multi-output Random Forest Classifier* proven effective in mapping the need for new formations in the simultaneous procurement of Civil Servant Candidates (CPNS). The use of computational models based on *machine learning* This model is capable of automating and synchronizing two main target variables: determining the minimum educational qualification threshold and estimating the financial compensation class. By utilizing real data from the 2024 national selection, the model has proven reliable in recognizing the characteristics and non-linear relationships of structured input variables, including agency name, position type, quota allocation, and geographic placement location.

In addition to predictive performance, the integration of methods *Mean Decrease Impurity* (MDI) as a metric *Feature Importance* successfully provides transparency into model decisions without getting bogged down in the characteristics *black-box* The results of the feature weighing analysis indicate that the parameters of agency name and position type are the most dominant factors influencing the final classification decision. The implications of these findings provide a strategic contribution for national personnel planners such as the National Civil Service Agency (BKN) and the Ministry of Administrative and Bureaucratic Reform (PANRB) to shift to a data-driven personnel management strategy (*data-driven decision-making*). The implementation of this model is recommended to mitigate the risk of placement inequality, eliminate bias in conventional job evaluations, and maintain the efficiency of government and the country's fiscal sustainability from inefficient employee spending budgets.

References

[1] B. State Civil Service, "State Civil Service Agency Strategic Plan 2020-2024,"*Journal of Bureaucratic Reform*, vol. 8, no. 2, pp. 45–52, 2020.

[2] R. M. Purnomo*et al.*, "Financial Impact of Inefficient Allocation of Employee Formation on Regional Budgets,"*Journal of Economics and Public Finance*, vol. 11, no. 3, pp. 112–120, 2022.

[3] S. Handayani, "Analysis of Civil Service Human Resource Needs Planning in Indonesia: Challenges and Solutions,"*Journal of Personnel Management*, vol. 9, no. 1, pp. 15–24, 2021.

[4] M. F. Azis and H. Kurniawan, "Accuracy of Mapping New Positions to Support the Effectiveness of the Wheels of Government,"*Journal of Public Administration*, vol. 14, no. 2, pp. 201–215, 2023.

[5] W. Utami*et al.*, "Job Class Prediction System and Remuneration Standards Using a Data-Based Comparative Approach,"*Indonesian Journal of Public Policy*, vol. 12, no. 4, pp. 330–338, 2024.

[6] A. A. Sutrisno*et al.*, "Analysis of Public Sector Workforce Qualification Needs Using Classification Algorithms,"*Journal of Information Technology and Systems Engineering*, vol. 10, no. 1, pp. 296–301, 2026.

[7] Regional Personnel Office, "Variations in Fiscal Capacity and Its Impact on Regional Employee Additional Income (TPP)"*Journal of Regional Autonomy*, vol. 7, no. 3, pp. 88–97, 2023.

[8] T. H. Prasetyo, "Application of Data Analytics in Synchronizing Central and Regional CPNS Formations,"*Scientific Journal of Governance*, vol. 15, no. 2, pp. 140–149, 2024.
[9] R. Saputra*et al.*, "Random Forest-Based Ensemble Model for Multi-Output Classification on Human Resource Management Data,"*Journal of Computer Science and Informatics*, vol. 6, no. 1, pp. 75–83, 2025.
[10] R. Saputra*et al.*, "Examining the Civil Servant Career Development Model Using Hybrid Particle Swarm Optimization-Random Forest Classification,"*Journal of Information Technology and Systems Engineering*, vol. 10, no. 2, pp. 310–316, 2026.
[11] T. H. Prasetyo, "Multi-Output Machine Learning Framework in Classifying Organizational Structure and Educational Qualifications of New Positions,"*Journal of Personnel Management*, vol. 14, no. 1, pp. 45–53, 2025.
[12] W. Utami*et al.*, "Development of an Intelligent System for Classifying Public Sector Remuneration Groups Based on an Optimized Gradient Boosting Machine Model,"*Journal of Public Policy and Finance*, vol. 11, no. 3, pp. 182–190, 2024.
[13] M. F. Azis and H. Kurniawan, "Performance Analysis of Multi-Class Random Forest Architecture on Historical Mapping of SSCASN Recruitment Data,"*Journal of Public Administration*, vol. 15, no. 2, pp. 201–209, 2024.
[14] S. Handayani*et al.*, "Multi-Stage Prediction Model for Determining Civil Servant Compensation Standards Using the Ensemble Tree-Based Feature Importance Approach,"*Journal of Computer Science and Informatics*, vol. 8, no. 4, pp. 122–130, 2025.
[15] A. A. Sutrisno, "Evaluation of the Random Forest Classifier Algorithm Based on Grid Search Cross-Validation for Public Manpower Needs Planning,"*Scientific Journal of Governance*, vol. 12, no. 1, pp. 88–96, 2026.