Development of Staff Evaluation Software Based on Association Matrix Methods and Data Mining Using the Streamlit Framework

Yosia Adi Susetyo[1], Hanna Arini Parhusip[2*], Suryasatriya Trihandaru [3]

[1,2,3,] *Master of Data Science, Faculty of Science and Mathematics, Satya Wacana Christian University, Indonesia*

*corr_author: hanna.parhusip@uksw.edu

This study discusses evaluating employee performance in microbiology laboratories using an association matrix implemented in web-based software with the Streamlit framework. The purpose of the research is to improve the employee performance evaluation process, which previously used conventional methods. This software is built from a sample receipt recording history data stored in a MySQL database. The initially unstructured data was processed using Python libraries such as NumPy, Matplotlib, Pandas, and Difflib to generate personnel evaluation information such as specialization, task duration, workload, and individual competencies. This software can provide a fast and accurate performance assessment according to the evaluation period. In a test with the System Usability Scale (SUS), the software scored 75.83, which was rated "good.". These results show that the software is easy to use and can improve the efficiency of employee performance evaluation. Follow-up tests with questionnaires given to 18 users showed that this system was preferable to previous conventional methods. This software helps laboratory managers evaluate employee performance effectively and efficiently.

Keywords: association matrix, data extraction, streamlit framework, personnel evaluation

I. INTRODUCTION

Performance measurement is an important part of any organization or corporate unit, as it serves as an instrument for achieving goals and targets. In this case, the microbiology laboratory, one of the working units in the company, must implement an effective and efficient working system to survive and continue to evolve. By understanding the performance of an employee or staff, we can assess the quality of that staff [1], [2]. In addition, performance assessments can be used for further company needs, such as recruitment considerations, assessments of organizational conditions, considerations of wage increases, and so on [3]. Performance assessment can also help management or supervisors diagnose and analyze organizational problems. This can be done by first understanding the symptoms and then starting to find a solution to the problem [4]. So far, the division of tasks has been based on supervisors' habits and manual notes, so the distribution of work is uneven. Each staff member has the same opportunity to manage all analysis parameters in distributing assignment letters. Each worksheet in the assignment letter presents a unique combination of work, necessitating adjustments and further processing the recording sample using Microsoft Excel. However, Microsoft Excel has limitations, especially in handling unstructured or textual data, as well as being unable to perform more in-depth data extraction. Furthermore, the company lacks specialized software incorporating evaluation features for calculating and measuring staff capacity.

Some studies related to staff assessment applications have been done by several authors which are done using an association matrix [5][6]. For instance, the association matrix is used to see patterns of plant-species interaction [7]. The utilization of local route planners for mobile robots that consider future obstacle positions. This research involves multi-resistance tracking, where the calculation of the association matrix is essential to ensure that the current frame cluster matches the previous frame clusters [8]. A web-based employee performance assessment application on PT.STARS, with a waterfall method system, can help to report employee data in each store more effectively and help the operational part track the overall progress of the store [9]. A website-based employee performance assessment application using the TOPSIS (Technique for Others Reference by Similarity to Ideal Solution) method at PT Petrochemistry Gersik [10]. Based on previous research above, the association matrix in staff performance assessment has not been used much for staff evaluation applications leading to the novelty of this research. Therefore, this study aims to determine the association matrix for distributing assignment letters fairly and build it into web-based staff evaluation software using the Streamlit framework. The laboratory can have the association matrix that illustrates the relationship between two variables and analysis parameters.

II. METHOD

A. Data Preparation

This step covers extraction, data exploration, and the implementation of association matrix models in web-based software, as presented in the process flow diagram shown in Fig. 1. Data was taken from the MySQL database by the fetching method. The data results from a digital recording of the web reception of samples in a microbiological laboratory that was previously recorded during the recording period from September 3, 2023, to February 1, 2024. Data consists of several features, such as an id, receiver_name, sample_type, parameter, analyst_name, sample_arrival, and sample_copletion [11]. There are about 13,000 input data points. The structure and features of the database tables are shown in Table I.

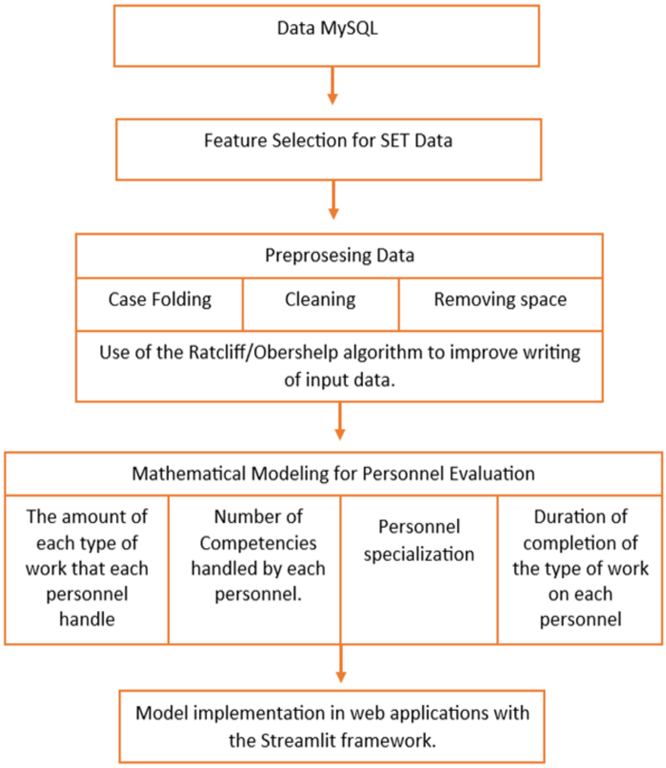

Fig. 1 Research methods and process flow

TABLE 1
STRUCTURE AND FEATURES OF THE MYSQL DATABASE

Name	Type	Collation	Null	Default
id	int(10)		No	None
receiver_name	varchar(30)	utf8mb4_general_ci	No	None
sample_type	varchar(50)	utf8mb4_general_ci	No	None
sample_name	varchar(500)	utf8mb4_general_ci	No	None
sample_quantity	float		No	None
parameter_quantity	float		No	None
parameter	varchar(50)	utf8mb4_general_ci	No	None
method	varchar(500)	utf8mb4_general_ci	No	None
analyst_name	varchar(25)	utf8mb4_general_ci	No	None
process_sample	varchar(30)	utf8mb4_general_ci	No	None
sample_arrival	date	utf8mb4_general_ci	No	None
sample_completion	date		Yes	NULL
unit_completion	date		Yes	NULL
report	date		Yes	NULL
source_code	varchar(500)	utf8mb4_general_ci	No	None
registration_number	varchar(50)	utf8mb4_general_ci	No	None
origin	varchar(30)	utf8mb4_general_ci	No	None
code_type	varchar(20)	utf8mb4_general_ci	No	None
last_editor	varchar(20)	utf8mb4_general_ci	No	None
created_at	datetime		No	None
modified_at	datetime		Yes	NULL
total_parameter	int(10)		No	None
total_days	int(10)		No	None

B. Preprocessing Data

The purpose of data preprocessing is to reduce data size, identify correlations or relationships between data, normalize data, eliminate outliers, and perform data transformation and reduction[12], [13]. Of the features in the database, only a few are taken for further processing, among them: analyst_name, parameter, sample_name, sampel_processing, sample_quantity, parameter_quantity, total_days, sample_arrival and sample_completion. The analyst_name feature provides information about the name of the staff who handle the assignment; the parameter gives information about what kind of jobs are contained in the assignments; sample_name provides information on the name of the sample being analyzed; sample_processing provides information about the status of the samples being worked on or completed; sample_arrival provides information about the date of acceptance of the task; and sample_completion information about assignments completed by the staff. The description of the data sets used is shown in Fig. 2. We focus on the "parameter" feature for data preprocessing. The first step is "case folding," i.e., all large characters are converted into small characters. The "cleansing" process serves to clear data from unnecessary read marks and symbols, then remove spaces aimed at merging words into one string [14]-[16].

C. Data Extraction

The lab has a list of the types of analysis obtained from analyzing data on a parameter feature. The analysis name is extracted based on the unique character in the parameter variable on the data set, and the result is shown in Fig. 3.

From the list of available job types, we implement the *Ratcliff/Obershelp* algorithm from the *Difflib* library and the *Sequence Matcher* function to correct input errors on the work data in the "parameter" feature. The system can automatically compare user input with the job list in the corpus and identify writing similarities [17], [18]. The system can improve the accuracy and consistency of the work data in the "parameter" feature. This process ensures that all entries reflect the correct standard name, as shown in Fig. 3. This will help reduce input errors and improve efficiency in data usage [19].

D. Mathematical Modeling for Evaluation

1) Evaluate the Amount of Work for Each Type of Work that Personnel Have Handled: Preprocessing data obtains the new data related to staff evaluation. The system will fill in data on an association matrix with a size of 12 (the number of staff) x 24 (the number of types of jobs present in the dataset). The size matrix increases when the system finds data on the parameter feature on each task sheet handled by a particular staff. For example, the assignment letter (called A) has the analysis parameter (work type) "alt, kapang, khamir", which is handled by the An5 staff. The matrix value (see Table 2) will be identified by value one on the column An5, the alt rows, kapang, and the khamir. Similarly, assignment letter B has the analysis parameters of salmonella, rhizobium, and kapang handled by the An1 staff, and its column is filled by one, where its row is kapang, rhizobium, and salmonella, respectively. Table II is a visual representation of the workload distribution among personnel for each analysis parameter, where the data on the matrix shows the individual contribution of each personnel.

Table II shows the calculation and filling of data on the matrix will run on the entire set of data, using (1).

$$J = \sum_{i=1}^{n} P_i \quad (1)$$

where *J* represents the total workload for each analysis parameter, *n* is the number of personnel involved, and P_i is the workload contribution handled by each *i*-th personnel. Equation (1) describes how the total workload is calculated by summing up the work done by each staff member in a single parameter.

2) The Amount of Competence Handled by Each Member of Staff: The amount of competence handled by each member of staff can be done by further processing the matrix of the number of types of jobs handled by each staff member by counting rows with nonzero values. Referring to Table II, the number of competencies that An1 possesses is three, i.e., kapang, rhizobium, salmonella, and An5 has the alt, the kapang, and the khamir competencies. The working system runs on all the datasets, so evaluations can occur during the selected period.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13377 entries, 0 to 13376
Data columns (total 10 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   analyst_name       13377 non-null  object
 1   parameter          13377 non-null  object
 2   sample_name        13377 non-null  object
 3   process_sample     13377 non-null  object
 4   sample_amount      13377 non-null  float64
 5   parameter_amount   13377 non-null  float64
 6   total_parameters   13377 non-null  int64
 7   days_count         13377 non-null  int64
 8   sample_arrival     13377 non-null  object
 9   sample_completion  11623 non-null  object
```

Fig. 2 Description of the datasets used in building a staff assessment system

```
[0] set()trichoderma
[1] {0}alt
[2] {0, 1}enterobacter
[3] {0, 1, 2}lactobacillus
[4] {0, 1, 2, 3}azotobacter
[5] {0, 1, 2, 3, 4}clostridium
[6] {0, 1, 2, 3, 4, 5}mpncoliform
[7] {0, 1, 2, 3, 4, 5, 6}azospirillum
[8] {0, 1, 2, 3, 4, 5, 6, 7}pseudomonas
[9] {0, 1, 2, 3, 4, 5, 6, 7, 8}shigella
[10] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}staphylococcusaureus
[11] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10}candidaalbicans
[12] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}bakteriasamlaktat
[13] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}bakteriselulolitik
[14] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13}biletolerantgramnegatifbacteria
[15] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14}totalcoliform
[16] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15}saccharomyces
[17] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16}khamir
[18] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17}aspergillus
[19] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18}pelarutphospat
[20] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19}salmonella
[21] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20}rhizobium
[22] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21}kapang
[23] {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22}ecoli
```

Fig. 3 List of available job types in the data set

TABLE II
SIMULATION OF WORKLOAD CALCULATIONS ON EACH STAFF

Parameter Analysis	Personnel Name											
	An0	An1	An2	An3	An4	An5	An6	An7	An8	An9	An10	An11
alt						1						
kapang		1				1						
khamir						1						
rhizobium		1										
salmonella		1										

3) Staff Specialization: Staff specialization is determined by calculating the number of staff with expertise in each type of job. If the number of staff skilled in a parameter is less than 3 people, then the staff who do the job is indicated as specialized or specialized personnel in dealing with the parameter.

4) The Duration of Completion of the Type of Work on Each Staff: The association matrix method is also used to determine how long each type of job is completed. This is done by computing the time each type of task is completed. Here is a simulation of the duration calculation of completion of the type of job in the assignment letter. For instance, the assignment letter A has analytical parameters = alt, kapang, and khamir that are handled by the An5 staff for 5 days. Task letter B has analytical parameters = alt, kapang, and rhizobium handled by An5 staff with a duration of 7 days. The simulation of filling matrix values is shown in Table III. Each assignment has a different duration of work. This work duration data is stored in the "total_days" feature, which is the difference between the date on the "sample_arrival" feature and the "sample_completion".

Based on the simulation in Table III, the average staff performance time on each parameter shows that the An5 staff took 6 working days to handle analytical parameters such as alt, 6 working days for the kapang, 5 working days for the khamir, and 7 working days in the rhizobium. The process carried out on the entire dataset on each assignment, as in the simulation, thus obtaining

TABLE III
SIMULATION OF WORK DURATION CALCULATION FOR EACH STAFF ON EACH ANALYTICAL PARAMETER

Parameter Analysis	Personnel Name											
	An0	An1	An2	An3	An4	An5	An6	An7	An8	An9	An10	An11
alt						= (5+7)/2						
kapang						= (5+7)/2						
khamir						= 5/1						
rhizobium						= 7/1						

results that can provide information about the competence of the staff in dealing with a particular type of job in this context is the speed of work during the specified evaluation period. The result of the calculation of the average speed of work on a particular staff and type of job can be mathematically formulated as (2).

$$R = \frac{1}{n}\sum_{i=1}^{n} T_i \qquad (2)$$

where R is the average speed of work on staff in dealing with each analytical parameter; n, represents the number of data rows that include completion times for staff and analytical parameters; T_i, represents completion time by staff *with* on a particular parameter. Equation (2) calculates the average time of work for each staff member involved in one parameter to determine the handling time of each type of analysis parameter and the result is shown in Table 3.

5) Software Development: We use Streamlit to develop our software. Streamlit is a *framework with* an open-source *Python programming language* used to build web-based applications to share large datasets [20][21]. In the software, we built algorithms of mathematical modeling embedded in the web, i.e., generating evaluation values for each staff [22]. The system will process and display the evaluation results.

6) Evaluating Implements System: SUS is a method used to assess the usefulness of a product. SUS consists of 10 questions, with the answer to each question being a score from 1 to 5 [23] [24]. Questions are presented in a Google form with the question types are identified as Q1, Q2, …, and Q10. Some questions are "I'm thinking of using this system again, I feel that this system is complicated to use, I find this system easy to use". The total score is calculated as follows: for each odd-numbered question, the score obtained from the user will be subtracted by 1. For each even-numbered question, the score obtained is 5 minus the user's score. The SUS score is obtained by adding all the scores from the 10 questions and multiplying the total by 2.5. The total score ranges from 0 to 100, and each score range has a rating category from A to F that determines the quality of the software being built [25]. The scoring index and SUS score are shown in Fig. 4.

7) Evaluation of System Performance Compared to Previous Evaluation Methods: This system was tested with a questionnaire on 18 users to assess the effectiveness of the system in obtaining information related to workload, the duration of personnel handling samples, evaluation of staff specialties, individual competencies, speed of information, and assessment accuracy. The results were compared to the manual calculation method and Excel calculation. The evaluation score had a score range of 1–5 points, which were classified as follows: 1 (poor), 2 (less), 3 (adequate), 4 (good), and 5 (excellent). The assessment results of 18 users will be averaged. The higher the average score of the method's assessment, the greater the user satisfaction. The test results will be presented in the form of a spider diagram to show the changes after the system has been implemented.

III. RESULT AND DISCUSSION

It has been shown the association matrix for staff evaluation, and then it was implemented on web-based software using Streamlit. The software can perform staff evaluations quickly and accurately. The user must log in, set the evaluation range, and press the data load button. After login input and data loading, the system will automatically display the required evaluation results. To access the staff evaluation page, the user must enter the employee's Perner number and password. Then, go to the staff assessment page and enter the evaluation period. If an evaluation of all the data is needed, just press the data load button, and the evaluation process will run. Software output results are described as follows.

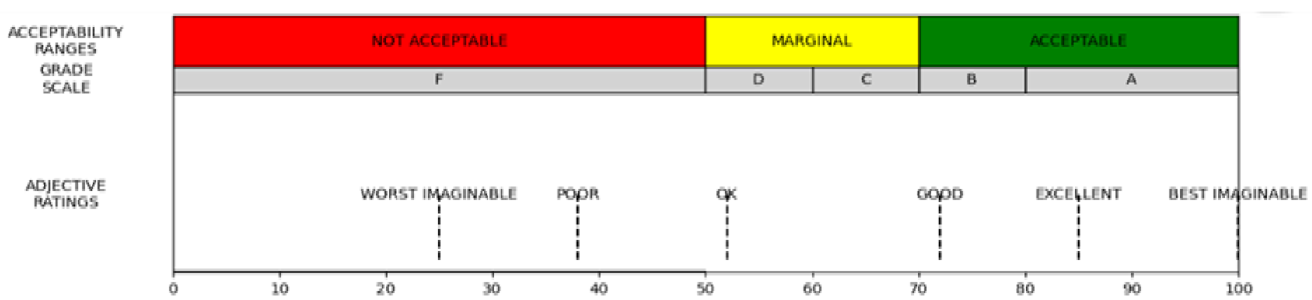

Fig. 4 SUS test scale on staff evaluation system

Table IV illustrates the number of microbiological analysis parameters performed during the observation period, with rows indicating the type of parameter and columns representing the initials of the personnel who performed them. Each cell contains a number that indicates the number of analyses that personnel have completed for the parameter type on the corresponding row.

Fig. 5 displays a bar chart illustrating the number of analysis parameters handled by each personnel (An0 to An11) for different analysis parameters. The horizontal axis depicts the personnel ID, while the vertical axis indicates the type of analysis parameter. Each color of the bar represents a different microorganism, suggesting that personnel manage different types of analysis. Personnel such as An2 and An6 handle significantly larger volumes of parameters, particularly for the azotobacter and khamir analysis parameters, in contrast to other personnel like An4 and An5, who have lower analytical responsibilities. In general, this graph shows that the analytical tasks that different employees are responsible for are very different. Some employees are more responsible for a lot of different types of analysis, which could mean that they are less productive or not given the right tasks within their teams.

Based on the software output, we can obtain information where not all staff have the same analytical abilities. After the evaluation of the results, there is a variation in the number of samples and parameters handled by each staff. Some staff handle many samples but are limited to certain parameters, while others handle fewer samples but can master almost all parameters. This phenomenon has become a major concern in managerial assessment to ensure that the workload is balanced. From the output processing results of the software, we can identify the analytics factually who have the highest, medium, and lowest workloads. This information could be responded to by further measures to ensure a more equitable distribution of the workload among staff.

The competence of each staff and the number of staff capable of dealing with a particular type of job. In Fig. 6 we depict the output given by the software listing type of microbes, number of analysts and number of competences of each staff and Fig. 7 shows its histogram.

TABLE IV
A PART OF THE ASSOCIATION MATRIX THAT SHOWS THE NUMBER OF TYPES OF JOBS HANDLED BY EACH STAFF

Type Job	An0	An1	An2	An3	…	An11
saccharomyces	0	0	0	15		0
trichoderma	0	0	0	5		0
bakteriselulolitik	0	0	5	5		0
salmonella	3720	2616	8044	3891		4802
...						.
alt	4474	6812	12826	5679		6130
aspergillus	1	0	0	15	…	0

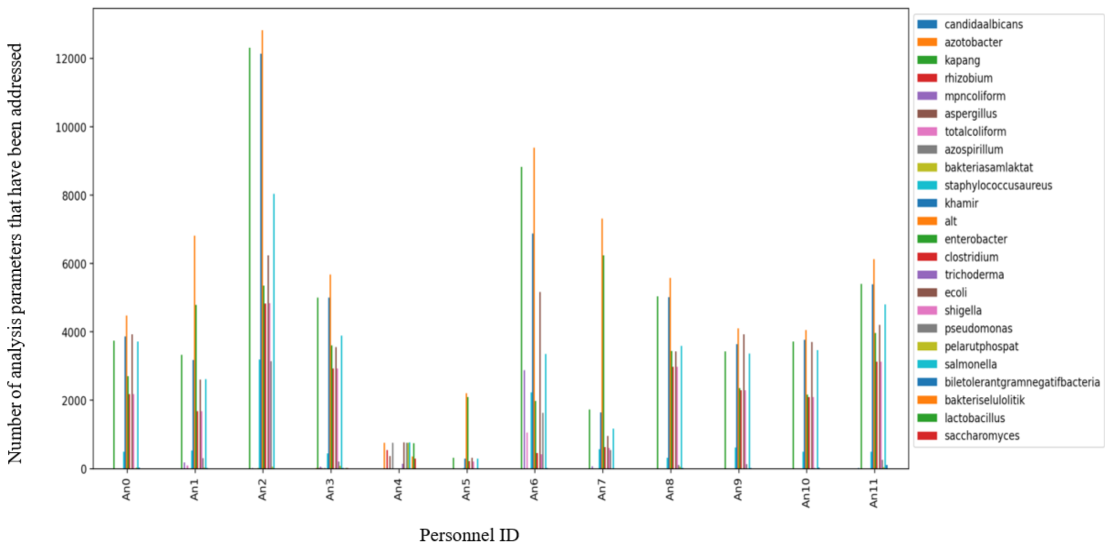

Fig. 5 The histogram of the workload handled by each

Fig. 6 shows the output generated by the software. There are kinds of jobs that can only be handled by two people. These observations can be the basis of the evaluation, where parameters that can only be dealt with by two analysts can expand their scope. As a precautionary measure, it is recommended to ensure that at least more than two people can handle each parameter. It could be an effective strategy to deal with situations where staff are unable to work for a certain period, thus ensuring the smooth operation of the laboratory. Fig. 7 displays the software's output, detailing the total competencies each staff member has mastered. We carry out competency measurement by calculating the number of analysis parameters handled during the research period. We consider each analysis parameter as having an equal contribution to competence, measuring the total competence based on the number of types of analysis parameters everyone has successfully managed. The total competencies directly reflect the quantity of parameter types handled during the period, as there is no difference in weights between the analysis parameters. Personnel Specialization in the context of this study, if the type of job is controlled only by less than or equal to two people, then further exploration will be made about who the staff are. By understanding who is responsible for the type of job, the supervisor, as an application user, can evaluate whether it is possible to transfer knowledge to other staff through specialized training or assignments [26].

Types of microbes	Number of Analysts	Types of microbes	Number of Analysts
rhizobium	4	azospirillum	5
lactobacillus	5	kapang	12
pelarutphospat	9	trichoderma	2
biletolerantgramne	8	candidaalbicans	2
enterobacter	11	pseudomonas	11
saccharomyces	3	khamir	12
shigella	11	salmonella	12
alt	12	azotobacter	5
bakteriasamlaktat	3	mpncoliform	8
clostridium	11	totalcoliform	4
staphylococcusaure	11	ecoli	12
aspergillus	4	bakteriselulolitik	3

Fig. 6 The distribution of personnel capabilities in each analytical parameter

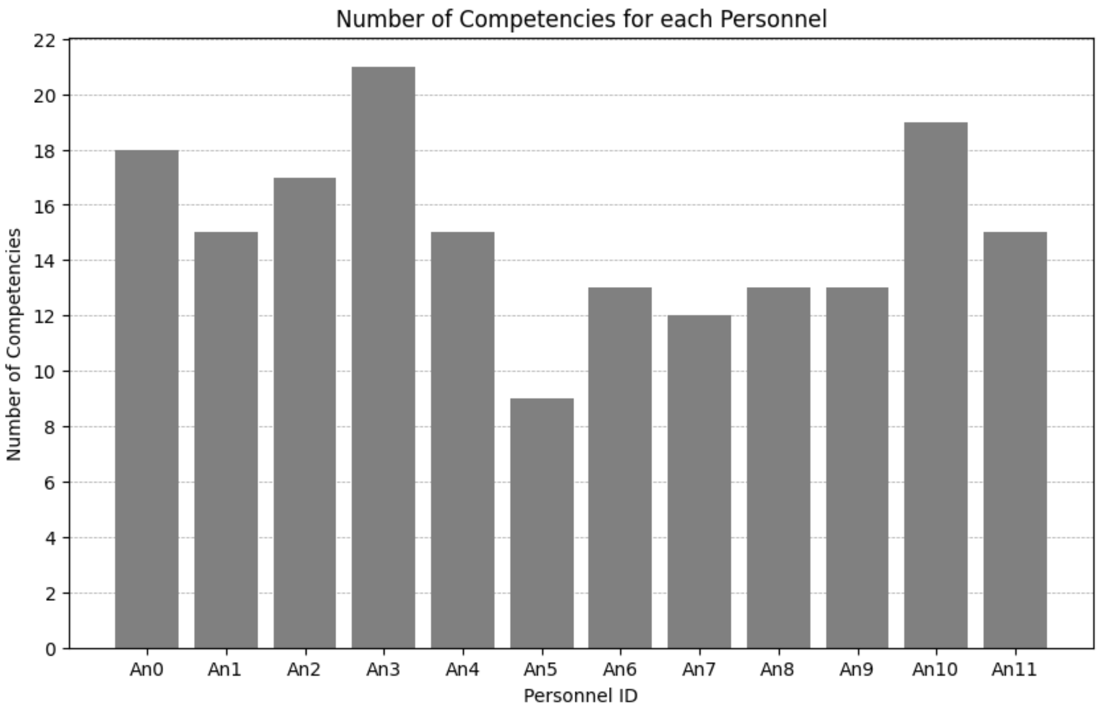

Fig. 7 The histogram of the number of types of work that each staff member can handle

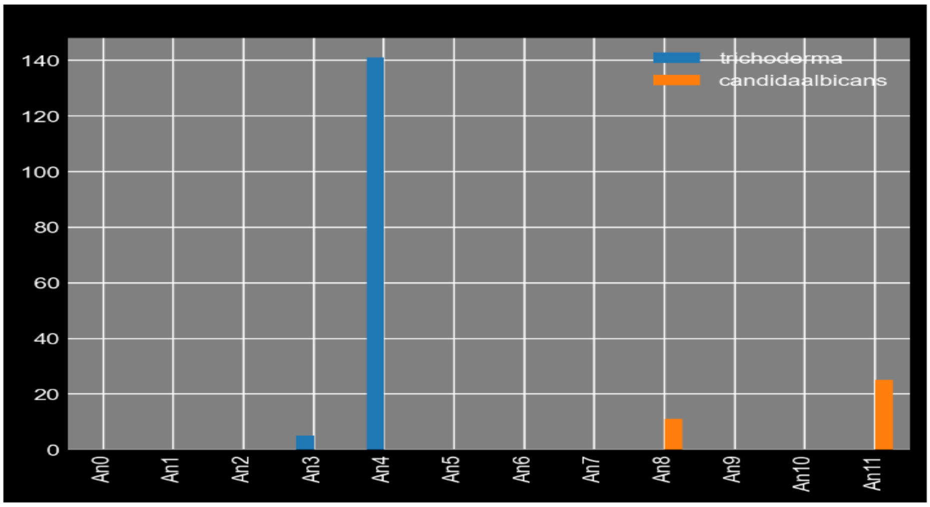

Fig. 8 System output in identifying personnel specialization

The identification process is also useful in understanding the competencies required to handle tasks based on analytical parameters. By defining the competencies required, the management team can plan appropriate training and skills development to ensure that the minimum number of staff to master each type of job is met [27]. The system works to produce outputs of personnel that belong to the specialization category within the specified observation period. The software output is shown in Fig. 7. Based on the output of the software, the tables and bar graph show that there are variations in the number of competencies each member of staff has. The total analysis parameters in the laboratory are 24. According to the software output, the information staff with the highest capabilities master only 21 parameters, while there are staff who master only 9 parameters. Evaluation can be done by improving the competence of staff with skills below average. It makes it possible to be an alternative when the number of incoming samples increases and the distribution of tasks is required. Improved competence can be achieved through interpersonal training, participation in seminars, and external training.

The output from the software also shows the number of people who can handle the kinds of work that are available in the lab. Based on these results, we can see the spread of the capacity of the staff to handle the type of work and the ability of the laboratory to analyze samples. The output from the software is shown in Fig. 6. Further analysis is carried out to see the type of work that can only be handled with the number of personnel less than or equal to two personnel with the software output shown in Fig. 8. During the study period, it was found that trichoderma and candidaalbicans were the types of analysis parameters that met these criteria. In-depth analysis is necessary to identify personnel who master these parameters as well as the number of samples they have handled, thus providing a clearer picture of their level of specialization. The graph displays the personnel name on the x-axis and the number of samples handled on the y-axis. The software's output results revealed the competence of two individuals in each parameter: trichoderma (An3 and An4) and candidaalbicans (An8 and An11). Management classifies these personnel as specialists who can train other staff. This software's output assists management in determining if other staff members require training to handle these parameters.

1) Sample Handling Speed

The software can generate output data that provides information on the time required by each person to complete each type of task in units of days. The matrix that associates the speed of handling the kind of work on each staff is shown in Table V and Fig. 9. to handle a particular analysis parameter. The values in the table indicate the number of days required for each personnel to handle the analysis parameters. A positive number signifies a measurable handling time, while a negative number (-1) indicates that the personnel has not been involved in handling or there is no data available. For instance, H_An0 took 10.4 days to treat salmonella, while H_An3 took 11.08 days. H_An0 required 12 days to handle aspergilus, while H_An3 required 7 days. The results of this analysis highlight the variation in handling among different personnel, reflecting the level of difficulty and the difference in efficiency in handling each type of parameter. Fig. 9 shows a bar graph where each bar represents one type of parameter analyzed by each personnel, with the time measured in days. The different colours on the bars represent different types of analysis parameters. Some of the observable patterns from this graph show that personnel have significant variations in handling time depending on the type of analysis parameters being handled. The number of bars in each personnel's name indicates the number of competencies they have mastered. In general, these charts make it simple to visualize the distribution and variation of working time among personnel for each analysis parameter, showing the level of efficiency or difficulty that personnel may experience in relation to a particular type of analysis.

2) The SUS test

The SUS test was conducted by giving an evaluation questionnaire on the software created to 18 application users. The 18 respondents were direct users of influential applications in determining policies, making assignment letters, and delegating tasks to each analyst. The respondents included managers, supervisors, admins, and assistant supervisors. As shown in Fig. 5, the ratings are set to Extraordinary , Good , OK, Bad, in the range of Above 80, Score70-80, 50-70, Below 50 ("Bad") respectively**.** The SUS tests showed a score of 75.83, with a "Good" rating, indicating that most users feel that the software has good usability and is well received. However, some areas may still be where minor improvements or adjustments are needed to improve the overall user experience. Fig. 10 displays the results of the assessment. Based on the test results, the average SUS score of 75.83 indicates a high level of system usability.

3) System Performance Evaluation

Software performance compared to previous methods, i.e., manual calculation, and the use of Excel, showed significant improvements in the staff evaluation process. Assessments include the effectiveness of systems in calculating staff workloads, personnel duration, searching for staff specialization, calculating individual competence, speed of information acquisition, and accuracy of assessments, with scores 4.6, 4.8, 4.0, 4.3, 4.8, 4.3 respectively. These results show that the software provides more dominant satisfaction values than previous conventional methods. A comparison of the assessment of the methods is shown in the spider diagram in Fig. 11.

TABLE V
SOME PART OF ASSOCIATION MATRIX SHOWING THE SPEED OF HANDLING THE KIND OF WORK ON EACH STAFF

Type Job	H_An0	H_An1	H_An2	H_An3	…	H_An11
saccharomyces	-1	-1	-1	7		-1
trichoderma	-1	-1	-1	7		-1
bakteriselulolitik	-1	-1	7	7		-1
salmonella	10.39962	8.228889	8.480339	11.0816		10.84056
clostridium	10.24684	8.549153	8.516456	11.22124		11.24124
...						.
alt	9.720126	7.736611	8.470859	10.74558		10.38883
aspergillus	12	-1	-1	7	…	-1

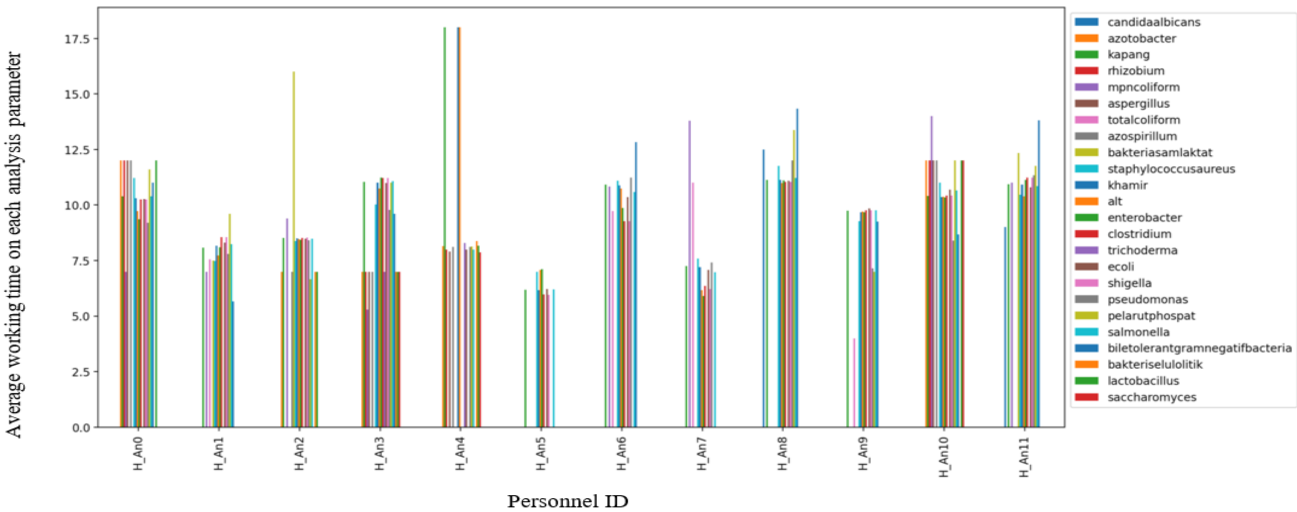

Fig. 9 Histogram the average velocity of every staff in handling each type of job

No	ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total Skor
1	90004546	3	3	3	4	2	3	3	3	3	2	72.50
2	90006264	3	3	2	3	3	3	3	3	1	0	60.00
3	90001457	3	4	3	3	3	3	3	4	1	1	70.00
4	90005644	4	3	3	3	3	3	2	2	2	2	67.50
5	90003090	4	3	3	3	3	2	2	3	3	1	67.50
6	90003746	4	4	4	4	2	3	3	4	3	3	85.00
7	90001330	4	2	3	2	3	2	2	2	2	2	60.00
8	90002083	4	3	3	2	4	3	3	3	4	1	75.00
9	90001627	3	4	3	2	3	2	3	3	3	1	67.50
10	90004306	3	3	3	3	2	3	2	3	3	3	70.00
11	90001737	4	3	4	2	3	2	3	3	2	1	67.50
12	90006462	4	3	4	3	3	4	3	3	3	4	85.00
13	90006171	3	4	3	4	3	3	3	3	3	4	82.50
14	90003035	4	4	4	3	4	4	4	4	4	4	97.50
15	90003048	3	4	4	3	3	4	3	4	3	3	85.00
16	90005524	4	3	3	3	4	3	3	3	3	2	77.50
17	90003092	4	4	3	4	3	3	4	3	3	2	82.50
18	90003601	4	4	4	4	4	4	2	4	3	4	92.50
Average SUS Score												75.83

Fig. 10 System Usability Scale (SUS) test results as website application evaluation

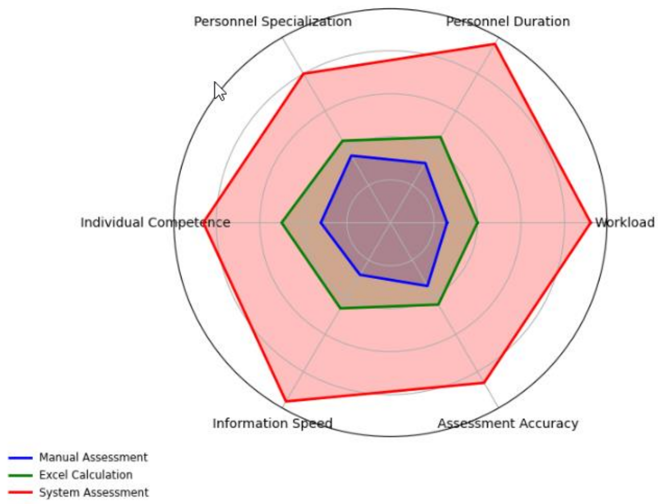

Fig. 11. A spider diagram of comparison of staff evaluation methods

The results show that staff evaluation methods using an association matrix approach and implemented in web systems have a significant impact compared to previously used staff assessment methods, which used manual calculations and Excel. By implementing this software, users can quickly and easily obtain the necessary information.

IV. CONCLUSION

This paper describes the association matrix to represent the task evaluation of personnel in a microbiology laboratory by using parameters such as workers' competencies and task assignments. Data extraction and exploration processes can be used to build web-based software with the Streamlit framework, which can enhance the process of personnel resource evaluation in the company's microbiology laboratory. This software output can be used to measure the workload, specialization, speed, and distribution of the ability of staff to deal with various work parameters. This information provides valuable insights for management in establishing measures for better organizational development. By implementing the software built, the evaluation process becomes easier because it only requires input from an evaluation period, and the system will deliver evaluation results quickly, accurately, effectively, and efficiently. The System Usability Scale (SUS) test results on the software showed a score of 75.83 with a "Good" status, indicating that the software was well received and used by users in microbiological laboratories. The results of the overall system performance testing were also carried out by comparing the evaluation process of staff using conventional methods and methods of using software. The results showed significant changes in management efficiency in calculating staff workloads, searching for staff specialization, calculating individual competence, speed of information acquisition, and accuracy of assessment.

ACKNOWLEDGEMENT

This research is internal research from Satya Wacana Christian University under the grant number 077/SPK-PF/RIK/08/2024.

REFERENCES

[1] D. Andreas, ‘Employee Perfomance: The Effect of Motivation and Job Satisfaction’, *J. Kepegawai. dan*

Organ., vol. 1, no. 1, pp. 28–35, 2022.

[2] A. Hartika, M. Fitridiani, and M. Asbari, 'The Effect of Job Satisfaction and Job Loyalty on Employee Performance: A Narrative Literature Review', *J. Inf. Syst. Manag.*, vol. 2, no. 3, pp. 9–15, 2023.

[3] A. Wilkinson, N. Bacon, T. Redman, and S. Snell, *The SAGE handbook of human resource management*, 12th ed. London: Kogan Page, 2010. doi: 10.4135/9780857021496.

[4] L. R. Sembiring, Y. R. Edward, and C. F. Rostina, 'The Effect of Work Motivation on Employee Performance with Job Satisfaction as an Intervening Variable at PT. Indonesia Tourism Development Corporation (ITDC)', *Undiksha J. Econ.*, vol. 10, no. 1, pp. 146 – 155, 2018.

[5] J. A. Mazzon, 'Using the methodological association matrix in marketing studies', *Rev. Bras. Mark.*, vol. 17, no. 5, pp. 747–770, 2018, doi: 10.5585/bjm.v17i5.4175.

[6] Y. Qin, Y. Peng, K. Huang, C. Zhou, and Y. Tian, 'Association Analysis-Based Cybersecurity Risk', *IEEE Syst. J.*, pp. 1–10, 2020, doi: 10.1109/JSYST.2020.3010977.

[7] Djufri, 'Determination of distribution pattern, association, and interaction of plant species particularly the grassland in Baluran National Park, East Java', *Biodiversitas J. Biol. Divers.*, vol. 3, no. 1, pp. 181–188, 2002, doi: 10.13057/biodiv/d030103.

[8] X. Ou, Z. You, and X. He, 'Local Path Planner for Mobile Robot Considering Future Positions of Obstacles', *Processes*, vol. 12, no. 5, 2024, doi: 10.3390/pr12050984.

[9] W. Mulyana, W. Putri Bunda, M. Rikky Syahputra, N. Maulana, T. Habincaran, and S. Wahyuni, 'Pengembangan Aplikasi Berbasis Web Untuk Penilaian Kinerja Karyawan Pada Pt. Stars Internasional', *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2318–2324, 2024, doi: 10.36040/jati.v8i2.9474.

[10] A. J. M. Novita, A. W. Rizqi, and H. Hidayat, 'Aplikasi Penilaian Kinerja Karyawan Berbasis Website Menggunakan Metode Topsis Di Pt Petrokimia Gresik', *JUSTI (Jurnal Sist. dan Tek. Ind.*, vol. 2, no. 3, p. 307, 2022, doi: 10.30587/justicb.v2i3.3913.

[11] H. A. Mumtahana, 'Optimization of Transaction Database Design with MySQL and MongoDB', *SinkrOn*, vol. 7, no. 3, pp. 883–890, 2022, doi: 10.33395/sinkron.v7i3.11528.

[12] S. K. R. Moosavi, A. Saadat, Z. Abaid, W. Ni, K. Li, and M. Guizani, 'Feature selection based on dataset variance optimization using Hybrid Sine Cosine – Firehawk Algorithm (HSCFHA)', *Futur. Gener. Comput. Syst.*, vol. 155, no. June 2023, pp. 272–286, 2024, doi: 10.1016/j.future.2024.02.017.

[13] S. A. Alasadi and W. S. Bhaya, 'Review of data preprocessing techniques in data mining', *J. Eng. Appl. Sci.*, vol. 12, no. 16, pp. 4102–4107, 2017, doi: 10.3923/jeasci.2017.4102.4107.

[14] K. Maharana, S. Mondal, and B. Nemade, 'A review: Data pre-processing and data augmentation techniques', *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.

[15] M. Z. Al-Taie, S. Kadry, and J. P. Lucas, 'Online data preprocessing: A case study approach', *Int. J. Electr. Comput. Eng.*, vol. 9, no. 4, pp. 2620–2626, 2019, doi: 10.11591/ijece.v9i4.pp2620-2626.

[16] M. Danubianu, 'Step By Step Data Preprocessing for Data Mining. A Case Study', *Int. Conf. Inf. Technol.*, no. 17-18 September 2014, pp. 1–8, 2015, [Online]. Available: https://www.researchgate.net/publication/288825433

[17] W. Hidayat, E. Utami, and A. D. Hartanto, 'Effect of Stemming Nazief Adriani on the Ratcliff/Obershelp algorithm in identifying level of similarity between slang and formal words', *2020 3rd Int. Conf. Inf. Commun. Technol. ICOIACT 2020*, pp. 22–27, 2020, doi: 10.1109/ICOIACT50329.2020.9331973.

[18] Y. Özçevik, F. Yücalar, and M. Demircioğlu, 'Determining a Proper Text Similarity Approach for Resume Parsing Process in a Digitized HR Software', *Celal Bayar Univ. J. Sci.*, vol. 18, no. 4, pp. 371–378, 2022, doi: 10.18466/cbayarfbe.1049845.

[19] B. Y. Ong, R. Wen, and A. N. Zhang, 'Data blending in manufacturing and supply chains', *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, pp. 3773–3778, 2016, doi: 10.1109/BigData.2016.7841047.

[20] J. M. Nápoles-Duarte, A. Biswas, M. I. Parker, J. P. Palomares-Baez, M. A. Chávez-Rojo, and L. M. Rodríguez-Valdez, 'Stmol: A component for building interactive molecular visualizations within streamlit web-applications', *Front. Mol. Biosci.*, vol. 9, no. September, pp. 1–10, 2022, doi: 10.3389/fmolb.2022.990846.

[21] Y. Cui, X. Shi, Y. Qin, Q. Wang, X. Cao, X. Che, Y. Pan, B. Wang, M. Lei, and Y. Liu, 'Establishment and validation of an interactive artificial intelligence platform to predict postoperative ambulatory status for patients with metastatic spinal disease: a multicenter analysis', *Int. J. Surg.*, vol. 110, no. 5, pp. 2738–2756, 2024, doi: 10.1097/JS9.0000000000001169.

[22] M. A. Sufian, W. Hamzi, S. Zaman, L. Alsadder, B. Hamzi, J. Varadarajan, and M. A. K. Azad, 'Enhancing Clinical Validation for Early Cardiovascular Disease Prediction through Simulation, AI, and Web Technology', *Diagnostics*, vol. 14, no. 12, pp. 1–44, 2024, doi: 10.3390/diagnostics14121308.

[23] M. Ollenschläger, F. Kluge, M. Müller-Schulz, R. Püllen, C. Möller, J. Klucken, and B. M. Eskofier, 'Wearable gait analysis systems: ready to be used by medical practitioners in geriatric wards?', *Eur. Geriatr. Med.*, vol. 13, no. 4, pp. 817–824, 2022, doi: 10.1007/s41999-022-00629-1.

[24] K. Boyd, R. Bond, A. Vertesi, H. Dogan, and J. Magee, 'How People Judge the Usability of a Desktop Graphic User Interface at Different Time Points: Is there Evidence for Memory Decay, Recall Bias or Temporal Bias?', *Interact. Comput.*, vol. 31, no. 2, pp. 221–230, 2019, doi: 10.1093/iwc/iwz019.

[25] S. Putri, Dwi Andini; Hadinata, Denda Rinaldi; Nurwahyuni, 'Website Usability Testing on the Information and Coordination Center of Covid-19 in West Java Using the System Usability Scale', *Jitk*, vol. 6, no. 1, pp. 3–8, 2020, doi: 10.33480/jitk.v6i1.1335.

[26] A. Jahmani, O. Jawabreh, E. A. AL Fahmawee, M. S. Almasarweh, and B. J. A. Ali, 'The Impact of Employee Management on Organizational Performance in Dubai's Five-Star Hotel Sector', *J. Stat. Appl. Probab.*, vol. 12, no. 2, pp. 395–404, 2023, doi: 10.18576/jsap/120206.

[27] J. A. Kuprenas, 'Implementation and performance of a matrix organization structure', *Int. J. Proj. Manag.*, vol. 21, no. 1, pp. 51–62, 2003, doi: 10.1016/S0263-7863(01)00065-5.