Intelligent Platform Using Natural Language Processing for Pre-Selection of Personnel Through Professional Values Required by Private Companies

José L. Silva[1][a], Vicente O. Moya[1][b], Daniel W. Burga[1][c] and Carlos A. Tello[2][d]
[1]*Faculty of Engineering, Peruvian University of Applied Sciences, Lima Metropolitana, Lima, Peru*
[2]*Department of Physics, Paulista State University, São Paulo, Brazil*

Keywords: Natural Language Processing (NLP), Professional Values, Cultural Alignment, Preselection Process, Intelligent Platform.

Abstract: The pre-selection process is essential for companies, as it ensures the recruitment of competent staff for each position, maintaining a positive working environment crucial to meeting organizational objectives. This research presents an intelligent platform for the pre-selection of personnel based on professional values. When the selection process is poorly executed, it can lead to economic and intangible losses, such as delays in project progress and team demotivation. The platform employs natural language processing (NLP) to analyse applicant data, making it easier to identify candidates that best suit the needs of the company. The results indicate that the intelligent platform achieves an 80% accuracy in its recommendations.

1 INTRODUCTION

As of 2022, more than 26 million video interviews and 5 million candidate evaluations have been conducted using artificial intelligence (Koutsoumpis et al.,2024). The COVID-19 pandemic has accelerated the shift from traditional face-to-face interviews to digital interviews. This situation has been crucial to increase the quality of human resources decisions in a company (Fernandes et al., 2021).

On the other hand, the use of machine learning and sentiment analysis in personnel selection open an opportunity to innovate in human resources management (Campion, 2024), allowing for more accurate decision-making even when the data set to be analysed is very broad (Radonjić, Duarte & Pereira, 2024). In the same vein, aspects such as data privacy and personnel selection biases caused by machine learning models without the sensitive ability to recognize human soft features have been questioned during early AI integrations within companies (Delecraz et al., 2022).

However, to support this research, the APA-AVI study demonstrated that the accuracy of personality trait assessments significantly improved when machine learning models were trained using observer-based reports (Koutsoumpis et al., 2024). This suggests that defining an organizational ideal of value—understood as the set of human qualities that align with the company's culture and make a candidate more likely to be hired—helps to reduce the high variance and bias often found in automated evaluation processes (Wang et al., 2024).

For the development of this analysis, some complementary points of view will be evaluated. On the one hand, manual coding, used to classify or analyse data in a traditional way, is expensive and, even with extensive training programs, high levels of reliability are not always achieved (Van Atteveldt et al. (2021). This highlights the limitations of traditional approaches to today's challenges. On the other hand, it proposes the use of advanced techniques, such as the KNN method and the

[a] https://orcid.org/0009-0002-5103-2881
[b] https://orcid.org/0009-0000-6429-6101
[c] https://orcid.org/0000-0003-0312-727X
[d] https://orcid.org/0000-0002-0369-8999

Silva, J. L., Moya, V. O., Burga, D. W. and Tello, C. A.
Intelligent Platform Using Natural Language Processing for Pre-Selection of Personnel Through Professional Values Required by Private Companies.
DOI: 10.5220/0013463000003964
In *Proceedings of the 20th International Conference on Software Technologies (ICSOFT 2025)*, pages 201-206
ISBN: 978-989-758-757-3; ISSN: 2184-2833

BERTEC model, to overcome these limitations and optimize the precision in the analysis (Latifi, Jannach& Ferraro, 2022). In the same vein, the use of pre-trained models has proven to be a successful strategy, one of these being BERT (Gu et al., 2022). Therefore, it is committed to perform a procedure based on BERT models, with the help of a data set of more than ten thousand records, it is expected that the training will exceed in F1 to 85% of adjustment, this basing the concept of ideal professional value on recruiters with years of experience and knowledge in the field (Cui, et al., 2021).

The structure of the research is organized as follows: Section II presents previous studies on Natural Language Processing (NLP) techniques applied to recruitment. Section III details the sources, variables and methods of data collection, including criteria for assessing the suitability of candidates for professional values. Section IV discusses the implementation of the PLN system for data processing and profile information extraction, using pre-trained fine-tuning models based on Wiki Large. Section V validates results using performance metrics, such as the MAE. It is necessary to show the analysis of graphical efficacy of the results (He et al., 2024). Finally, section VI summarizes the main findings, limitations and possible future improvements to the system.

2 RELATED WORKS

Regarding the country of origin of this research, 3 out of 9 sources come from the Americas region. There are studies in which methodologies of perception of personality traits and the relationship to certain behaviors were evaluated. It is also proposed that the way candidates express themselves can significantly affect the outcome of a job interview, especially when using AI tools (Martín-Raugh et al., 2023). Under the same study parameters, the impact of correct personnel selection was evaluated using PLN and XGBoost techniques, whose central study process involved the collection of approximately 1.2 million job reviews (Feng, 2023). It was concluded that a 1% increase in review ratings is correlated with an increase from 0.68% to 0.73% in the company's market value (Feng, 2023).

Regarding research techniques, natural language processing (NLP) has been applied to tourism through sentiment analysis, tokenization and lemmatization, techniques that could be adapted in human resources to identify professional values in applicants.

In fact, the NLP achieved accuracy rates above 80%, suggesting that, in staff selection, it could improve the identification of candidates aligned with organizational culture (Koutsoumpis et al., 2024).

There is also a common approach at the intersection between PLN technology and human evaluation (Campion & Campion, 2024; Delecraz et al., 2022; Koutsoumpis et al., 2024). For example, regression analyses are used to determine the impact of vocal and visual characteristics on hiring decisions, finding a significant correlation with an effect size of 0.40 (Delecraz et al., 2022).

On the other hand, the readability and clarity in the texts generated by artificial intelligence has been evaluated, noting that models such as ChatGPT and Bard achieved almost identical relevance scores, with averages of 4.93 and 4.92, respectively (Campion & Campion, 2024). These studies illustrate how predictive models and artificial intelligence-driven analysis contribute to decision-making in Human Resources, sentiment assessment, and performance forecasting (Campion & Campion, 2024; Delecraz et al., 2022; Koutsoumpis et al., 2024).

In this sense, the implementation of algorithms based on machine learning improves accuracy and equity in identifying suitable candidates for different positions (Campion & Campion, 2024; Delecraz et al., 2022; Koutsoumpis et al., 2024).

3 METHOD

The developed platform introduces an innovative approach to assessing professional values and soft skills by combining natural language processing (NLP) and deep learning techniques. Using embeds, the platform transforms a set of key organizational values into vector representations, efficiently stored in pickling files. Similarly, it converts more than 10,000 candidate descriptions into embeds that allow for high-precision comparisons. Subsequently, the results that identify the most relevant values per candidate are consolidated into an Excel file, optimizing the flow of data to a Fine-Tuning process with BERT Distil. This trained model integrates advanced contextual learning and tuning techniques, providing a detailed and scalable assessment within the final platform.

3.1 Data Source Collection

The main data source for this investigation comes from a private company that provided a dataset of 19,555 applicant records, with confidential information such as names, contact details and censored addresses. These data include detailed curriculum information, descriptions and internal performance appraisals. The registers cover both structured data (age, gender, position applied for) and unstructured data (description of skills, professional achievements, personal description, cultural compatibility).All personal information was anonymised prior to processing, in compliance with the company’s internal data protection protocols and applicable privacy regulations.

3.2 Data Collection

The collected data were pre-processed using tokenization, lemmatization and sentiment analysis techniques. These processes ensure the standardization and cleanliness of texts, as demonstrated in previous research on NLP applied to Human Resources (Álvarez-Carmona et al., 2022).

Table 1: Related characteristics.

Characteristics	Values
Cultural Compatibility	0-100%
Model Accuracy Reached	<82%
NLP model	BERT, DistilBERT
Number of training records	19,555
Number of values trained	10
Main skills shown	3

3.3 Modular Procedure

Data processing was carried out using previously trained NLP models, BERT and DistilBERT, adjusted to assess professional values and cultural compatibility. These models were selected for their effectiveness in reducing biases and their ability to perform in-depth analyses in complex texts (Gu et al., 2022; Wang et al., 2024).

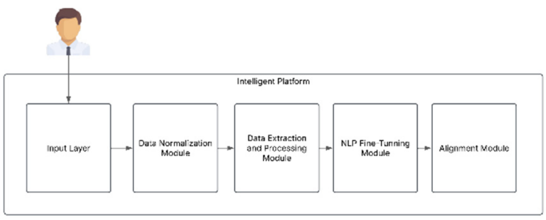

Figure 1: Modular structure of the intelligent platform.

The platform operates through five sequential modules:

- Input Layer:Itreceivestwo inputs—candidate CVs and recruiterselected values.
- Data Normalization Module: Responsible for loading, reading, and preprocessing the input data to ensure consistency.
- DataExtraction and Processing Module: Text data is tokenized, embeddings are generated for both CVs and values, and alignment scores are computed. Results are saved in a Pickle file.
- NLP and Fine-tuning Module: Uses the Pickle data to apply a fine-tuned DistilBERT model for refined matching.
- Alignment Module:Calculates alignment percentages and ranks candidates by best fit.

The platform was developed in Python, using Transformers (Hugging Face) for model fine-tuning, scikit-learn for evaluation, Pandas for data handling, and PyTorch for embedding generation and similarity computation.

4 EXPERIMENTAL CONTEXTS

The data was collected by the company over an estimated period of 2 years and 9 months, from October 2021 to July 2024, resulting in a total of 421 recruitment campaigns with more than 19,555 candidates for various positions.

Table 2: Number of applicants per year.

Year	Number of applicants
2021	3170
2022	8335
2023	6940
2024	1110

Table 3: General basics of training.

Code	AABAMA
Professional Summary	Computer Engineering student at PUCP, with an interest in video games, data analysis, business, software, and UX/UI. Advanced English proficiency, proactive, and motivated.
Response 1	I want to apply data analysis to make better decisions in business.
Response 2	I have learned to balance studies and personal life, making key decisions during difficult times.
Response 3	I wish to improve in data analysis to transform information into valuable insights.
Response 4	Yes, I want to pursue a master's degree in data analysis.
Response 5	Is there another way to access Support benefits if I do not obtain the position?

Each record stores information about a candidate, and it is essential that the experimental dataset focus primarily on columns containing descriptive information and candidate self-perception data.

4.1 Data Preparation

In this stage of experimentation and construction of the predictive model, it is important to consider that the questions posed by the evaluating authority are treated as predictive elements and have weight in the adjustment phase. The responses of each applicant (records) are based on the assessment elements shown in the following table.

Table 4: Impact of predictive elements on the model.

Predictive elements	**Weight %**
Do you plan to study for a master's degree or a doctorate?	12.5
Professional summary	12.5
What do you expect from your future career?	12.5
What are your long-term goals?	12.5
What have been its main achievements and challenges?	12.5
What have been your main interests?	12.5
What topics or types of work do you want to focus on in the future?	12.5
What questions would you like to ask us?	12.5

4.2 Training Model

This approach proposes a system that generates vector representations (embeddings) for both values and applicant descriptions. A model is then fine-tuned to predict the values that best correspond to a specific description.

4.2.1 Embedding Generation for Values and Descriptions

The embedding generation process converts both professional values and applicant descriptions into n-dimensional vectors in a vector space. This allows for measuring similarity based on their proximity in that space. For a given value $\mathbf{v_i}$ and description $\mathbf{d}$, the embedding is defined as:

$$e_v = embedding(v_i) \in \mathbb{R}^n \tag{1}$$

4.2.2 Measuring Similarity Between Descriptions and Values

To determine which values are most aligned with the applicant's description, we calculate the cosine similarity between the embeddings of the description and each value. Cosine similarity is defined as:

$$sim(e_d, e_v) = \frac{e_d \cdot \mathrm{e_v}}{||e_d|| \cdot \mathrm{e_v}||} \tag{2}$$

This formula measures the angle between the two vectors $\mathbf{e_d}$ and $\mathbf{e_v}$, where a value close to 1 indicates a higher similarity between the description and the value, while a value close to -1 indicates they are completely different.

4.2.3 Model Fine-Tuning

The next step is to fine-tune a pre-trained model to learn how to predict values more accurately, using tagged data to correct their errors. The optimized loss function during fine-tuning is the cross entropy, defined as:

$$L(\hat{y}, y) = -\sum_{i=1}^{V} y_i \log(\hat{y}_i) \tag{3}$$

In this case, $\mathbf{y^i}$ is the probability that the model predicts for value $\mathbf{V_i}$, and $\mathbf{y_i}$is the true label indicating if the value is relevant to the applicant. The loss function is minimized using gradient descent, adjusting the model weights $\boldsymbol{\theta}$ as follows:

$$\theta \leftarrow \theta - \eta \nabla_\theta L(f_\theta(e_d), y) \tag{4}$$

4.2.4 Predicting Relevant Values

Once adjusted, the model can predict the most relevant values for a new description of the applicant. To achieve this, compare the similarity between embedding the new description and embedding all values, selecting the closest matches.

$$\hat{v} = argmax_{v_i} sim(e_{d_{new}}, e_{v_i}) \tag{5}$$

This approach converts values and descriptions into comparable numerical representations, optimizing the model with Fine-Tuning to predict values more accurately. The use of F1 scoring has been shown to be highly effective, especially for predicting multiple relevant values at once. The effectiveness of the model will be observed after the combination of both techniques.

5 RESULTS

Table 4 shows that as the number of training samples increases, both ROC AUC and F1 Score improve consistently, reaching up to 0.9661 and 0.9482 respectively in the final iteration. Training time

Table 5: Summary of fit and accuracy metrics.

Step	Roc Auc	F1	Runtime	Samples per second
800	0.88163	0.816044	26.4994	222 005
1600	0.929502	0.889046	26.6936	220.39
2400	0.933901	0.901786	26,2108	224.45
3200	0.947302	0.916705	26.6532	220,724
4000	0.950816	0.920298	26.3827	222,987
4800	0.954999	0.928422	26.5894	221,253
5600	0.956941	0.930526	27.9548	210.447
6400	0.961154	0.937032	26.8393	219 193
7200	0.961192	0.934783	27.5994	213.157
8000	0.961487	0.939702	26.9416	218.361
8800	0.964093	0.941079	26.7634	219.815
9600	0.959429	0.935578	26.9127	218,571
10400	0.964106	0.942293	26.9424	218.354
12000	0.965494	0.94506	26.8664	218,972
19555	0.966073	0.948213	27.4084	214,642

remains stable (~27 seconds), confirming processing efficiency. Notably, gains start to stabilise after around 8800 samples, suggesting diminishing returns beyond that point.

Overall, DistilBERT fine-tuning outperformed the previous RoBERTa-based model (96% vs. 84% accuracy), offering greater precision in identifying candidates aligned with organisational values.

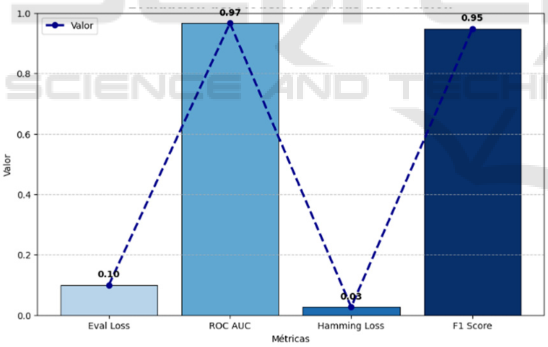

Figure 2: Metrics evaluation and comparison.

This figure illustrates the comparative performance of the proposed DistilBERT model across four key evaluation metrics: Eval Loss, ROC AUC, Hamming Loss, and F1 Score. The results show that the model achieves a low evaluation loss and Hamming Loss, both under 0.1, while obtaining high ROC AUC and F1 Score values, both exceeding 0.94. This balance indicates strong predictive capacity with minimal error, confirming that the model not only makes accurate predictions but also maintains consistency across multiple classification labels. The values were normalised between 0 and 1 for comparative clarity.

As for the impact on the pre-selection process, the implementation of DistilBERT led to a 12% reduction in selection errors, optimizing the recruitment process by minimizing incorrect decisions. This reduction represents a significant improvement in platform accuracy.

Table 6: Validation and Comparison with Terman Software.

	Terman					**IntelligentPlatform**				
ID	**Vr**	**Nr**	**Me**	**Rm**	**Ca**	**Ex**	**Pd**	**In**	**Int**	**Ta (%)**
1	75	70	65	72	68	85	88	78	88	84.6
2	60	68	72	65	70	75	65	75	80	75
3	55	50	60	55	60	78	82	85	82	76.3
4	88	85	90	85	88	80	78	76	84	80.6
5	60	55	58	65	68	88	90	82	90	84.4
6	45	40	50	48	52	72	68	60	70	69.8
7	78	75	80	77	80	82	85	85	86	81
8	50	55	60	58	65	78	82	85	84	81.6
9	82	80	85	80	82	85	84	80	85	81.4
10	65	60	62	67	68	80	85	78	88	83

Table 7: Meaning of abbreviations.

Abbreviation	**Meaning**
Vr	Verbal Reasoning
Nr	Numerical Reasoning
Me	Short-term memory or retention capacity
Rm	Mechanical Reasoning
Ca	Abstract Comprehension or Analytical Capacity
Ex	Excellence
Pd	Professional Development
In	Innovation
Int	Integrity
Ta (%)	Total alignment %

This table presents a comparison between Terman Test results and the intelligent platform's inferred professional values. While the Terman dimensions focus on cognitive abilities, the platform evaluates alignment with organisational values such as Excellence, Integrity, and Innovation. The alignment percentage reflects the consistency between both approaches, with most candidates scoring above 75%. This suggests that the platform is capable of approximating traditional human assessment criteria through automated value analysis, offering a scalable alternative for early-stage candidate evaluation.

Training time with DistilBERT was reduced by 72%—from 6 hours to just 1 hour and 40 minutes—without compromising performance. This efficiency, combined with high accuracy, enhances the platform's capacity to accelerate recruitment while ensuring alignment with company values and supporting more objective decision-making in HR.

6 CONCLUSIONS

The developed platform applies NLP and deep learning models to analyse and select candidates aligned with organizational values, achieving 96% accuracy and a 12% reduction in selection errors. This system not only optimizes the recruitment process by identifying more compatible profiles, but also improves cultural integration and objectivity in hiring decisions, allowing the company to build more cohesive teams aligned with strategic objectives.

REFERENCES

Álvarez-Carmona, M., Aranda, R., Rodríguez-Gonzalez, A. Y., Fajardo-Delgado, D., Sánchez, M. G., Pérez-Espinosa, H., Martínez-Miranda, J., Guerrero-Rodríguez, R., Bustio-Martínez, L., & Díaz-Pacheco, Á. (2022). Natural language processing applied to tourism research: A systematic review and future research directions. Journal of King Saud University - Computer and Information Sciences, 34(10), 10125–10144. https://doi.org/10.1016/j.jksuci.2022.10.010

Campion, E. D., & Campion, M. A. (2024). Impact of machine learning on personnel selection. Organizational Dynamics, 53(1), 101035. https://doi.org/10.1016/J.ORGDYN.2024.101035

Cui, Y., Che, W., Liu, T., Qin, B., & Yang, Z. (2021). Pre-Training with Whole Word Masking for Chinese BERT. IEEE/ACM Transactions on Audio Speech and Language Processing, 29, 3504–3514. https://doi.org/10.1109/TASLP.2021.3124365

Delecraz, S., Eltarr, L., Becuwe, M., Bouxin, H., Boutin, N., &Oullier, O. (2022). Responsible Artificial Intelligence in Human Resources Technology: An innovative inclusive and fair design matching algorithm for job recruitment purposes. Journal of Responsible Technology, 11. https://doi.org/10.1016/j.jrt.2022.100041

Feng, S. (2023). Job satisfaction, management sentiment, and financial performance: Text analysis with job reviews from indeed.com. International Journal of Information Management Data Insights, 3(1). https://doi.org/10.1016/j.jjimei.2023.100155

Fernandes, T., São, H., Pereira, J., & Pereira, V. (2023). Artificial intelligence applied to potential assessment and talent identification in an organizational context. Information Sciences, 2. https://doi.org/10.1016/j.heliyon.2023.e14694

Gu, Y., Tinn, R., Cheng, H., Lucas, M., Usuyama, N., Liu, X., Naumann, T., Gao, J., & Poon, H. (2022). Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. ACM Transactions on Computing for Healthcare, 3(1). https://doi.org/10.1145/3458754

He, W., Farrahi, K., Chen, B., Peng, B., & Villavicencio, A. (2024). Representation transfer and data cleaning in multi-views for text simplification. Pattern Recognition Letters, 177, 40–46. https://doi.org/10.1016/J.PATREC.2023.11.011

Huang, X., Yang, F., Zheng, J., Feng, C., & Zhang, L. (2023). Personalized human resource management via HR analytics and artificial intelligence: Theory and implications. Asia Pacific Management Review, 28(4), 598–610. https://doi.org/10.1016/j.apmrv.2023.04.004

Koutsoumpis, A., Ghassemi, S., Oostrom, J. K., Holtrop, D., van Breda, W., Zhang, T., & de Vries, R. E. (2024). Beyond traditional interviews: Psychometric analysis of asynchronous video interviews for personality and interview performance evaluation using machine learning. Computers in Human Behavior, 154. https://doi.org/10.1016/j.chb.2023.108128

Latifi, S., Jannach, D., & Ferraro, A. (2022). Sequential recommendation: A study on transformers, nearest neighbors and sampled metrics. Information Sciences, 609, 660–678. https://doi.org/10.1016/j.ins.2022.07.079

Latifi, S., Mauro, N., &Jannach, D. (2021). Session-aware recommendation: A surprising quest for the state-of-the-art. Information Sciences, 573, 291–315. https://doi.org/10.1016/j.ins.2021.05.048

Martin-Raugh, M. P., Leong, C. W., Roohr, K. C., & Chen, X. (2023). Perceived conscientiousness and openness to experience through the relationship between vocal and visual features and hiring decision in the interview. Computers in Human Behavior Reports, 10. https://doi.org/10.1016/j.chbr.2023.100272

Raman, R., Venugopalan, M., & Kamal, A. (2024). Evaluating human resources management literacy: A performance analysis of ChatGPT and bard. Heliyon, 10(5). https://doi.org/10.1016/j.heliyon.2024.e27026

Radonjić, A., Duarte, H., & Pereira, N. (2024). Artificial intelligence and HRM: HR managers' perspective on decisiveness and challenges. European Management Journal, 42(1), 57–66. https://doi.org/10.1016/J.EMJ.2022.07.001

van Atteveldt, W., van der Velden, M. A. C. G., &Boukes, M. (2021). The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms. Communication Methods and Measures, 15(2), 121-140. https://doi.org/10.1080/19312458.2020.1869198

Wang, L., Zhou, Y., Sanders, K., Marler, J. H., & Zou, Y. (2024). Determinants of effective HR analytics Implementation: An In-Depth review and a dynamic framework for future research. Journal of Business Research, 170, 114312. https://doi.org/10.1016/J.JBUSRES.2023.114312.