Neural Computing and Applications
https://doi.org/10.1007/s00521-024-10587-3

S.I.: MULTIMODAL DATA FUSION AND MACHINE LEARNING FOR SMART CITIES

Hyper-personalized employment in urban hubs: multimodal fusion architectures for personality-based job matching

Dipika Jain[1] · Saurabh Raj Sangwan[2] · Akshi Kumar[3]

Received: 1 November 2023 / Accepted: 1 October 2024

Abstract

In the evolving landscape of smart cities, employment strategies have been steering towards a more personalized approach, aiming to enhance job satisfaction and boost economic efficiency. This paper explores an advanced solution by integrating multimodal deep learning to create a hyper-personalized job matching system based on individual personality traits. We employed the First Impressions V2 dataset, a comprehensive collection encompassing various data modalities suitable for extracting personality insights. Among various architectures tested, the fusion of XceptionResNet with BERT emerged as the most promising, delivering unparalleled results. The combined model achieved an accuracy of 92.12%, an R2 score of 54.49%, a mean squared error of 0.0098, and a root mean squared error of 0.0992. These empirical findings demonstrate the effectiveness of the XceptionResNet + BERT in mapping personality traits, paving the way for an innovative, and efficient approach to job matching in urban environments. This work has the potential to revolutionize recruitment strategies in smart cities, ensuring placements that are not only skill-aligned but also personality-congruent, optimizing both individual satisfaction and organizational productivity. A set of theoretical case studies in technology, banking, healthcare, and retail sectors within smart cities illustrate how the model could optimize both individual satisfaction and organizational productivity.

Keywords Multimodal deep learning · Smart cities · Personality detection · ChaLearn dataset · Urban employment strategies

1 Introduction

As we step into the embrace of smart cities, we're concurrently stepping into the age of the Fourth Industrial Revolution—a revolution that heralds the convergence of digital, biological, and physical innovations [1]. In this vibrant juncture, traditional employment paradigms are being rigorously challenged, as the nexus of technology and urban sophistication necessitates a fresh perspective on workforce placement. Traditionally, employment strategies, especially in bustling urban environments, prioritized tangible metrics such as qualifications, years of experience, and technical proficiencies [2]. Such metrics, while imperative, often failed to capture the comprehensive essence of a candidate. To illustrate, consider two candidates, both with stellar resumes showcasing impeccable technical skills and identical years of experience. Yet, beneath this surface uniformity lie stark differences in their personalities, work preferences, and interpersonal aptitudes. One might be an extrovert, thriving in collaborative scenarios, while the other might excel in solitary, deep analytical tasks. Traditional hiring metrics may overlook these disparities, but in the nuanced professional environments of today's urban hubs, such distinctions could be critical.

Modern urban hubs, with their diverse populations and equally diverse professional landscapes, offer vast opportunities. Yet, these opportunities come intertwined with challenges. In the quest to match the right individual with

✉ Akshi Kumar
Akshi.Kumar@gold.ac.uk

1 Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

2 School of Computer Science and Engineering, Galgotias University, Greater Noida, India

3 Department of Computing, Goldsmiths, University of London, London, UK

the right job role, cities and corporations have started realizing that the 'right match' goes beyond just aligning skills. Individual personalities, encompassing traits, preferences, motivations, and behaviours, play a pivotal role. For instance, an extroverted individual with a penchant for social interactions might excel in client-facing roles or team leadership positions. In contrast, an introverted individual, who finds solace in solitude and deep work, might be better suited for research or individual contributor roles.

As urban companies continue their relentless evolution, there is a palpable shift towards recognizing the importance of individual personalities in employment. Such hyper-personalized approaches discern that job satisfaction and optimal performance are not solely contingent on skill alignment but require a deeper resonance between job roles and individual personalities. To bring this concept closer to reality, consider the vibrant tech sector in cities like San Francisco or Seoul. Here, tech enterprises, ranging from fledgling startups to established giants, often grapple with finding the 'right fit' for their teams. A conventional hiring approach might yield technically proficient candidates, but a personality misalignment could culminate in reduced job satisfaction, and consequentially, diminished productivity. However, by integrating advanced profiling that goes beyond skills to encompass personality traits—like innovative thinking, communication styles, or adaptability—the hiring process can evolve. For instance, a role demanding constant stakeholder interaction might be best suited for someone with an extroverted, communicative personality, even if their technical prowess is slightly inferior to a more introverted counterpart. Further, imagine the potential of an AI-driven tool that can analyse a candidate's LinkedIn articles to assess their industry insights, their Twitter feed to gauge their communication style and topical interests, and their participation in online forums to understand their problem-solving approach. Such a tool, when fine-tuned ethically and with respect to privacy concerns, could provide recruiters with a goldmine of information, enabling hyper-personalized job matches that account for personality traits just as much as, if not more than, skills.

To realize the vision of hyper-personalized employment and achieve a comprehensive understanding of personality traits, an advanced methodology is paramount—one that can combine and interpret data from diverse sources. Guided by this objective, we turned to the First Impressions V2 dataset [3]. This dataset, with its expansive collection of 10,000 video clips and meticulously curated transcripts, stands as an exemplary source of multimodal data. Originating from a broad spectrum of YouTube HD videos, these clips provide more than just visual indicators. Averaging a duration of 15 s, the clips are classified according to the five-factor model, assessing personality traits such as Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. They offer insight into vocal nuances, non-verbal behaviours, and various facets of human communication. Concurrently, the transcribed text presents a rich analysis of linguistic patterns, vocabulary preferences, and other linguistic markers indicative of personality attributes. The dataset has been crucial in exploring the relationship between audio-visual cues and personality perceptions.

Our primary objective is to develop a nuanced understanding of personality traits through multimodal data analysis and to apply this understanding in creating a sophisticated job matching system. This system is designed to align individual personality traits with the optimal job roles, enhancing both employee satisfaction and organizational productivity. Employing a comprehensive array of data processing techniques, including video and transcript analysis through state-of-the-art convolutional neural networks [4] and transformer architectures [5] like XceptionResNet and BERT, our methodological approach is robust and innovative. We focus on extracting and interpreting nuanced personality cues from video clips and textual data, aiming to provide a holistic personality assessment. In this research, a systematic approach was undertaken to provide an enriched understanding of video and transcript processing, as well as the models applied. Summary of the research steps undertaken is as follows:

(a) Video Processing

- The videos were pre-processed to extract 3 frames per video at equal intervals and audio in the WAV format.
- Audio data were transformed into images representing Mel spectrogram, chromagram, and spectral bandwidth, suitable for CNNs.
- Common audio and image processing tasks included noise reduction, resizing, cropping, etc., using libraries like OpenCV, PIL, and librosa.

(b) Transcript Processing

- Transcriptions were pre-processed using Tokenizers.
- They were cleaned to remove punctuations and irrelevant tokens.
- The processed transcripts were tokenized using BertTokenizer and DebertaTokenizer.

(c) Models Used

- For video (frames and audio as images) modality: InceptionV3, EfficientResNet, and XceptionResNet were used.
- For text modality (transcripts): BERT and DeBERTa were used.
- For multimodal data, fusion models were created combining features of InceptionV3, Efficient ResNet, and XceptionResNet with BERT.

Such a multifaceted approach, augmented by AI and machine learning, can revolutionize urban employment. Envision an intelligent system capable of analysing a candidate's public presentations, gauging their expertise, their communication styles on professional networks, and even their problem-solving discourse on platforms like GitHub or Stack Overflow. The significance of this study lies in its potential to revolutionize recruitment strategies, enabling a deep resonance between an individual's personality and their professional role. By advancing the capabilities of AI in personality assessment, we pave the way for more effective and satisfying employment in the fast-evolving urban job market.

The rest of the paper is organized as follows: firstly, the paper talks about the shift from skill-based to personality-driven approach, further section discusses the related work followed by the detailed description of the dataset. The methodology section, that is Sect. 4 of the paper, is emphasized on the deep neural architecture and the multimodal fusion models which further relates to the analysis of the results obtained with graphical representations in Sect. 5.

2 The urban shift: from skill-based to personality-driven employment

Modern urban centres are not just places of work; they are melting pots of diverse cultures, backgrounds, and personalities. Companies in these hubs are often involved in multifaceted projects, requiring teams with varied skills and personalities. Here is where the results of personality trait detection become pivotal. Take the tech industry, for example. While a developer's technical knowledge is undeniably crucial, their personality traits like adaptability, resilience, or inclination for teamwork can significantly influence their overall contribution to a project. A startup environment, characterized by rapid changes and multitasking, might prioritize adaptability and openness over other traits. In contrast, a well-established tech giant, with its structured workflows, might value consistency and detail-orientation.

Moreover, urban hubs are increasingly becoming the centres for global businesses, meaning employees often collaborate with teams across different parts of the world. Here, cultural sensitivity, a component of agreeableness, becomes indispensable. An individual's ability to understand, appreciate, and work harmoniously with colleagues from diverse backgrounds can be a game-changer. This is where the results from personality trait detection techniques shine. By analysing vast datasets, such as video interviews, we can derive accurate personality profiles of candidates. These profiles, when matched against a role's requirements, can result in better hiring decisions. For example, roles that require constant stakeholder interaction might prioritize extraversion, while positions involving deep research might value introversion and conscientiousness. Further, modern businesses aren't static; they evolve, pivot, and adapt. An employee's ability to grow with the company, learning new skills and adapting to new roles, is often influenced by their personality. Traits like openness to experience and adaptability can be strong indicators of an individual's potential for growth and versatility.

However, this shift towards personality-driven employment does not come without challenges. The foremost concern is ensuring that the process remains unbiased. Relying heavily on personality assessments might inadvertently lead to a monoculture, where only certain personality types are favoured, leading to a lack of diversity in thought and perspective. Moreover, while personality traits are crucial, they cannot completely replace skill assessments. A balance must be struck where skills and personality are both given their due importance. Over-relying on one at the expense of the other can lead to suboptimal hiring decisions.

The urban shift from skill-based to personality-driven employment is a testament to the complexities of modern work environments. As urban hubs continue to grow, in both size and diversity, understanding an individual's personality will become even more crucial. However, the future likely lies in a balanced approach, where skills and personality are both factored into hiring decisions. Personality trait detection offers a powerful tool in this paradigm, enabling companies to make more informed, holistic hiring choices. In conclusion, as urban hubs redefine employment, integrating the insights from personality trait detection can lead to more harmonious, productive workplaces. While the road ahead might have its challenges, the potential rewards—in terms of employee satisfaction, productivity, and business growth—are immense.

2.1 Impact of personality impressions on interview outcomes in personality-driven employment

The impact of personality impressions on interviews is crucial in the modern job market, especially in a personality-driven employment landscape. These impressions provide valuable insights that help employers make strategic hiring decisions that benefit both the individual and the organization in the long term. Personality traits not only influence how candidates are perceived by potential employers but also determine how well they fit into the company's culture and the specific job role. Here are the key points highlighting this importance:

- *First Impressions* Interviews are often the first point of contact between a candidate and an employer, where initial impressions are formed. Personality impressions during these interactions can greatly influence an interviewer's perception of a candidate's suitability for the role, often affecting the final hiring decision.
- *Cultural fit* Organizations look for individuals who not only have the necessary skills and qualifications but also fit well with their organizational culture. Personality impressions help employers assess whether candidates' values, work ethic, and behaviour align with the company's ethos.
- *Predicting job performance* Employers often use personality impressions as indicators of future job performance. Traits like conscientiousness, agreeableness, and emotional stability are linked to positive work outcomes and can be pivotal in predicting a candidate's success in the role.
- *Enhancing team dynamics* The personality impression can provide insights into how a candidate might interact with potential colleagues and contribute to team dynamics. Employers aim to select individuals whose personality traits will complement and enhance the existing team structure.
- *Reducing turnover* By assessing personality impressions, employers can make more informed hiring decisions that lead to longer-lasting employment relationships. Hiring individuals whose personalities are a good fit for the job and company can reduce turnover rates and associated costs.

3 Related work

Over the past decade, several studies have delved into the realm of personality trait recognition using various data modalities, such as text, audio, and visual inputs. In 2008, Pianesi et al. [6] conducted pioneering research on this topic, applying machine learning techniques to understand personality traits, including the Big Five model, during social interactions. This work has had significant implications across psychology, human–computer interaction, and the social sciences. Subsequent study in 2014 by Alam and Riccardi [7], explored the integration of multimodal data for personality analysis, utilizing advanced machine learning approaches.

In more recent years, research has focused on enhancing the interpretability and practical applications of personality analysis. In 2018, Escalante et al. [8] introduced an innovative method for modelling and explaining apparent personality traits using video data, increasing accuracy and transparency. In the same year, Gorbova et al. [9] synergized visual and textual data to improve personality trait analysis, demonstrating its superiority over baseline methods. In 2019, Junior et al. [10] conducted a comprehensive survey of vision-based techniques for evaluating apparent personality traits, consolidating existing knowledge and highlighting practical applications in human–computer interaction and social robotics. These studies collectively contribute to our understanding of how diverse data sources can be harnessed to gain insights into human behaviour and personality traits within social contexts, with wide-ranging implications in various fields.

Additionally, in 2019, Wingate et al. [11] explored how individuals form initial impressions, especially in situations involving deception or "faking." The study examined the attributions people make when encountering deceptive behaviour and discussed the practical implications of these early impressions in contexts like hiring decisions, interpersonal relationships, and legal proceedings. In 2020, Nørskov [12] and colleagues investigated the dynamics of human–robot interaction in job interviews, shedding light on the ethical concerns and psychological implications associated with robotic interviewers. Finally, in 2021, Aslan et al. [13] and Song et al. [14] continued to advance the field by combining models for improved accuracy and exploring self-supervised learning for automated personality trait recognition based on facial dynamics, highlighting practical applications in human–computer interaction and affective computing.

4 Methodology

In this research, a systematic approach was undertaken to provide an enriched understanding of video and transcript processing, as well as the models applied. The following subsections discuss the details of dataset and techniques used.

0	He's cutting it and then turn around and see t...	0.523364	0.488889	0.626374	0.552083	0.601942
1	Responsibility to house the organ I had been g...	0.345794	0.366667	0.472527	0.375000	0.582524
2	I actually got quite a few sets of black pens ...	0.252336	0.511111	0.406593	0.291667	0.485437
3	I ate a lot. I'd like a lot of foods. I rememb...	0.457944	0.377778	0.505495	0.489583	0.398058
4	Now I'll ask you guys to leave a question in t...	0.607477	0.622222	0.406593	0.489583	0.621359
...	...	...	...	...	...	...
9995	Du du du. No, that's not why I made this video...	0.289720	0.300000	0.208791	0.312500	0.135922
9996	They do it all the time.	0.719626	0.722222	0.670330	0.781250	0.572816
9997	Comfortable and I don't want anyone to feel un...	0.355140	0.677778	0.472527	0.395833	0.446602
9998	You're not giving yourself enough calories to ...	0.467290	0.622222	0.527473	0.645833	0.669903
9999	Eat enough carbs, eat more fats to get in more...	0.654206	0.588889	0.813187	0.635417	0.728155

Fig. 1 Transcripts of the dataset

4.1 The First Impression dataset

First Impression ChaLearn is a publicly available largest Big-5 personality dataset. It constitutes 10,000 videos labelled into various types of sentiments displayed by them. Using the videos, we have generated novel datasets of various modalities like frames, audios, and text.

The videos have an average duration of 15s, extracted from more than 3,000 HD YouTube videos of people speaking in English while facing the camera. Each video has five labels in the range [0,1] which depict the Big Five personality traits, namely—Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness. The videos are also accompanied by their transcripts, in the form of a pickle file. The transcript contains 435,984 words, averaging to 43 words per video, with a maximum of 85 words in a single transcript and 14,535 unique words (Figs. 1 and 2).

4.2 Pre-processing of video and textual modalities

For video processing, the initial steps involved the extraction of three frames from each video at consistent intervals, coupled with the retrieval of audio in the WAV format (Fig. 3). These audio data were subsequently converted into image formats, encapsulating representations like Mel spectrogram (Fig. 4), chromagram (Fig. 5), and spectral bandwidth (Fig. 6), making them apt for analysis via convolutional neural networks (CNNs). To ensure the accuracy and clarity of these image and audio outputs, a series of processing tasks were employed, such as noise

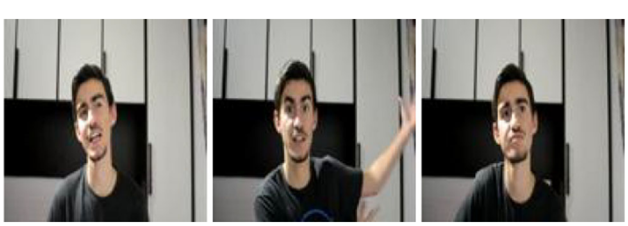

Fig. 2 Frames of the dataset

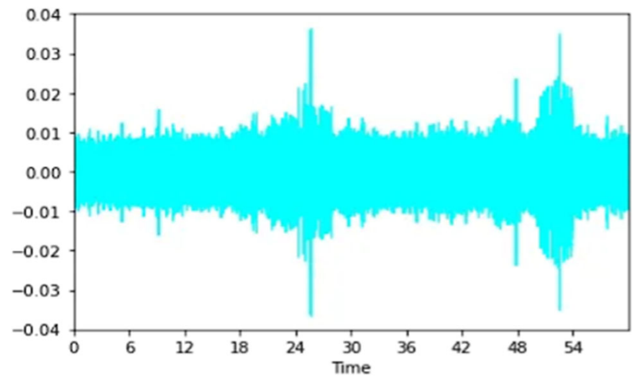

Fig. 3 Audio data in WAV format

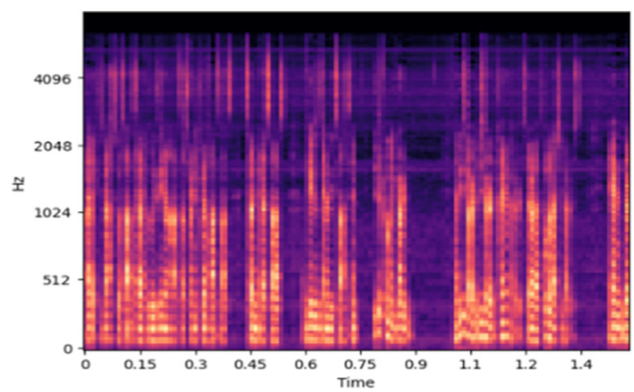

Fig. 4 Mel spectrogram

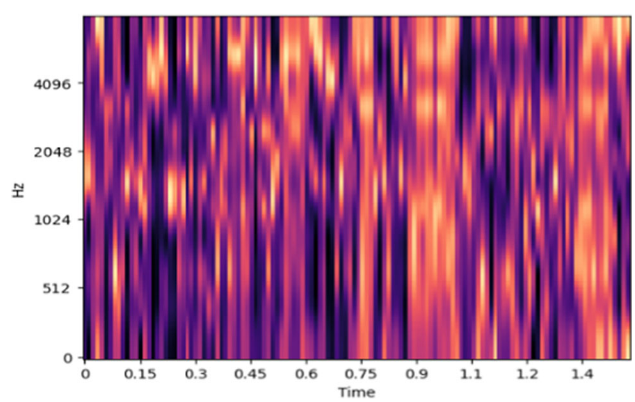

Fig. 5 Chromagram

reduction, resizing, and cropping. Renowned libraries like OpenCV, PIL, and librosa proved instrumental in this phase. A pre-processed audio data are shown in Fig. 7.

Transitioning to transcript processing, transcriptions of the videos were provided in pickle files with the annotations in the dataset itself and thus used directly for the models. The transcript data have been pre-processed using Tokenizers after filtering out punctuations and irrelevant tokens. The transcript contains 435,984 words, averaging to 43 words per video, with a maximum of 85 words in a

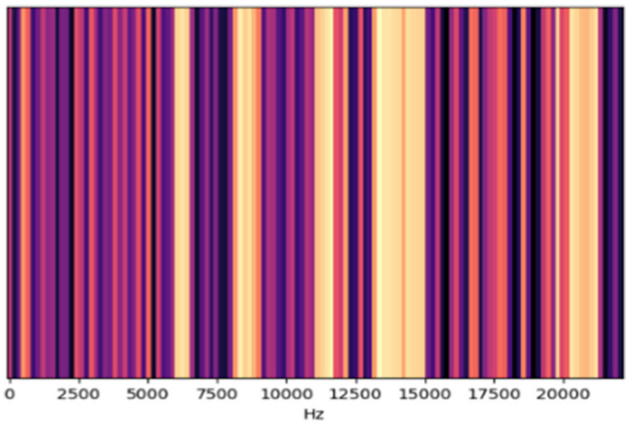

Fig. 6 Spectral bandwidth

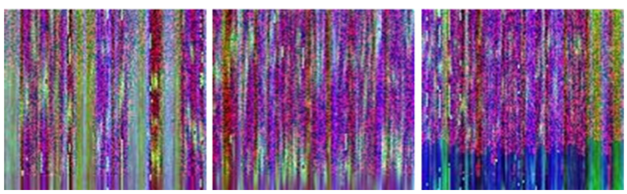

Fig. 7 (128,128,3) shaped audio images generated after pre-processing

single transcript and 14,535 unique words. Using the transcription and annotation pickle files, we have generated a CSV file which included a text transcript and the 5 labels of Big Five traits corresponding to the videos the transcript has been extracted from. The data-frame generated from the csv file is shown in Fig. 1. These data are tokenized using BertTokenizer and DebertaTokenizer, respectively, for both the models, after adding [CLS], [SEP] and [PAD] tokens to it.

4.3 Deep neural architectures

When it came to model application, a varied yet targeted approach was adopted. The video modality, which encompasses both frames and the transformed audio images, was analysed using models such as InceptionV3 [15], EfficientResNet [16], and XceptionResNet [17]. In contrast, the textual content derived from transcripts was put through transformer models like BERT [18] and DeBERTa [19]. Our recent paper [20] demonstrates the applicability of transformer-based models in understanding complex personality traits from textual data. Recognizing the potency of a combined approach, fusion models were developed, seamlessly integrating features from video-focused models with the textual capabilities of BERT, ensuring a comprehensive multimodal data analysis. With our processed video frames and audio feature images in hand, the next crucial step is data fusion. We fuse the processed video frames and audio feature images into a single input tensor. This integration of modalities allows our model to learn from both visual and auditory cues simultaneously.

4.3.1 InceptionV3

InceptionV3, an advanced convolutional neural network introduced by Google, is designed to improve computational efficiency and accuracy. It utilizes inception modules with various filter sizes (1×1, 3×3, 5×5) and pooling operations to efficiently capture features at different scales. The factorized 7×7 convolutions further reduce computational complexity, allowing for better performance with fewer parameters compared to traditional architectures. InceptionV3 has demonstrated remarkable results in computer vision tasks like image classification and object detection. Its versatility makes it suitable for various applications, and pre-trained models can be fine-tuned for specific tasks through transfer learning. However, InceptionV3's drawback lies in its high computational cost and memory requirements, making it challenging for resource-constrained environments. Additionally, its architecture might not be as effective for tasks involving extensive sequential data modelling, such as certain natural language processing tasks.

4.3.2 EfficientResNet

EfficientNet and ResNet are two prominent convolutional neural network architectures known for their performance and efficiency. EfficientNet scales width, depth, and resolution uniformly, balancing accuracy and computational cost. ResNet addresses deep network training issues with its residual or skip connections, allowing for deeper and more effective training. We've merged both architectures into EfficientResNet, a model that incorporates ResNet's skip connections into the EfficientNet design. The fused data, which now include information from both video frames and audio features, serve as the input to our EfficientResNet hybrid model. The final output is regression-based.

4.3.3 XceptionResNet

Xception (Extreme Inception) is a deep convolutional neural network architecture that builds upon the idea of the Inception architecture. It replaces the standard convolutional layers in Inception modules with depth-wise separable convolutions, which significantly reduces the number of parameters and computations while maintaining or even improving performance. This architecture aims to capture complex patterns in a more efficient manner. ResNet

(Residual Network), on the other hand, introduces residual connections that allow gradients to flow more easily during training, enabling the training of very deep networks. Residual connections involve adding the original input to the output of a layer, which helps in alleviating the vanishing gradient problem. This innovation has enabled the creation of extremely deep neural networks that were previously difficult to train. We have developed an innovative architecture called XceptionResNet, harnessing the collective strengths of Xception and ResNet designs to engineer a potent and resource-efficient neural network. By leveraging the individual advantages of Xception and ResNet paradigms, this hybridized approach engenders a unified model marked by heightened performance and computational efficiency.

4.3.4 BERT

BERT (Bidirectional Encoder Representations from Transformers) is a transformer-based neural network architecture that revolutionized natural language processing (NLP). Introduced in 2018, BERT addresses the contextual understanding challenge in NLP by training on a large corpus of text data using a masked language modelling objective. This allows BERT to capture deep contextual relationships between words, making it highly effective for a wide range of NLP tasks. The architecture of BERT consists of multiple layers of self-attention mechanisms, enabling it to capture intricate dependencies between words in both directions. By leveraging BERT's bidirectional contextual understanding, the processed transcripts can be analysed in-depth to discern subtle nuances in language use, emotional expressions, and linguistic patterns. This bidirectional approach empowers the model to capture not only what is explicitly stated but also the underlying context and relationships between words, allowing for a more comprehensive and accurate assessment of individuals' personality traits, including the Big Five (Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism). BERT's advanced natural language processing capabilities make it a valuable tool for uncovering the intricacies of human communication, enabling insights into the rich tapestry of personalities expressed through written or spoken language.

4.3.5 DeBERTa

DeBERTa (Decoding-enhanced BERT with Disentangled Attention) is an influential advancement in the realm of natural language processing (NLP), building upon the foundation of BERT. Introduced as A model architecture to address limitations in understanding global context, DeBERTa leverages disentangled self-attention mechanisms to enhance contextual comprehension. By untangling different aspects of language, DeBERTa excels in capturing long-range dependencies, making it well-suited for tasks requiring extensive sequential data modelling.

4.4 Multimodal fusion architectures

This section presents a comprehensive overview of the fusion models employed in personality trait prediction. In the preceding section, the deep neural architectures described are now employed in an integrated manner, incorporating the technique of weighted fusion. Weighted fusion is a valuable approach applied to amalgamate data from diverse modalities within the realm of multimodal data analysis. This technique grants the capability to allocate differing levels of significance to the contributions of each modality, contingent upon their pertinence and effectiveness in a given task. By doing so, it enhances the overall performance and resilience of multimodal systems, ensuring that the most pertinent and impactful information from various sources is effectively harnessed for comprehensive analysis and prediction. The first architecture, Inception V3 + BERT, combines InceptionV3 and BERT models to simultaneously process textual, image, and audio data. The video modality is processed using InceptionV3, while the text modality undergoes transformation by the BERT transformer. The concatenated outputs from these models pass through a series of densely connected layers to predict the Big Five traits. Similarly, the EfficientResNet + BERT architecture combines EfficientResNet and BERT for multimodal data processing. EfficientResNet handles image inputs, while BERT processes text inputs. These outputs are fused using a weighted fusion technique and subsequently processed through dense layers to predict personality traits. The final fusion architecture, XceptionResNet + BERT, utilizes Xception blocks with residual connections for image inputs and pretrained BERT for text data, ultimately employing weighted fusion and dense layers to predict personality traits. Figure 8 illustrates the XceptionResNet + BERT multimodal fusion architecture.

Incorporating findings from the First Impression dataset used in this study, we can leverage advanced AI models to decode subtle cues in candidates' expressions, speech patterns, and behaviours that signal underlying personality traits. This method enables a nuanced understanding of how a candidate's personality aligns with the cultural and functional demands of a job role. By applying a model trained on this dataset, recruiters can systematically evaluate and match candidates to positions where they are most likely to thrive, thereby optimizing job satisfaction and organizational productivity. This approach not only streamlines the recruitment process but also ensures a

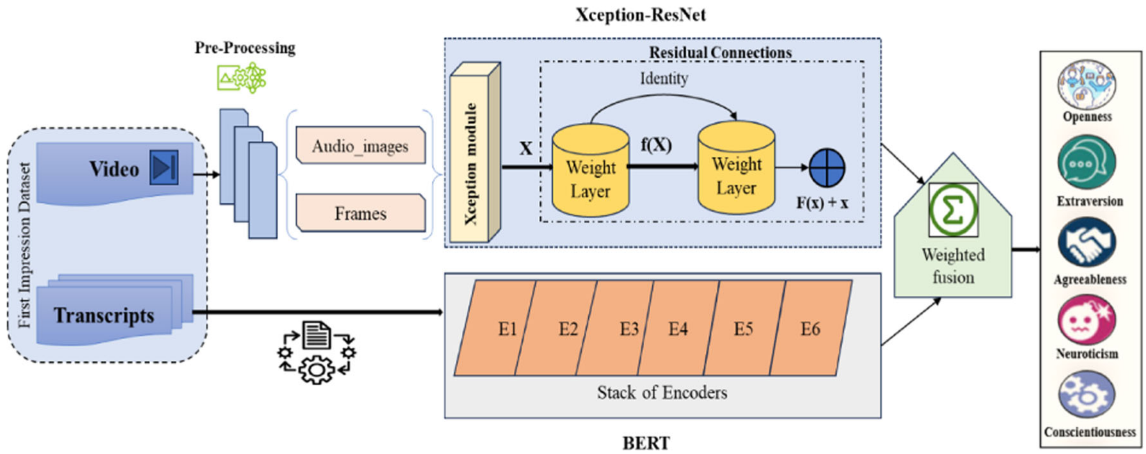

Fig. 8 XceptionResNet + BERT structure

deeper harmony between an individual's innate characteristics and their professional responsibilities, embodying the essence of personality-driven employment.

5 Results and discussions

This section presents a comprehensive analysis of the experimental results obtained from the multimodal models used for predicting personality traits. Evaluation metrics such as accuracy, F1-score, precision, and recall are employed to assess the performance of these models. The discussion highlights the model performance across different data modalities and provides insights into the effectiveness of each approach. The following hyperparameters (Table 1) have been tuned while training the models:

The following metrics have been used while analysing the models [21]:

- *Accuracy* The degree to which our predicted outputs conform with the true labels of the data. It has been calculated by first calculating the mean difference between the predictions and target labels, and subtracting it from 1, giving us a measure of how close our predictions were to the true labels.
- *R2 score* The R2 score, or coefficient of determination, measures the proportion of the variance in the dependent variable that is explained by the independent variables in a regression model. It is calculated by dividing the explained variance by the total variance and ranges from 0 to 1, where a higher value indicates a better fit of the model to the data.
- *Mean squared error (MSE)* It quantifies the average squared difference between predicted and actual values in a regression model, providing a measure of the model's prediction accuracy.

Table 1 Model hyperparameters

Model	Hyperparameter
BERT	Learning Rate = 5e-5
DeBERTa	Learning Rate = 1e-4
InceptionV3	Learning rate = 0.03
	Alternate ReLU and Tanh activation functions for Dense layers
	SGD optimizer with 0.9 momentum instead of Adam allows the gradient to not get stuck at a local minimum easily
EfficientResNet	Learning rate = 0.01 initially, with an exponential decay with weight decay = 0.3 over 500 steps (approximately 2.5 epochs)
	Sigmoid dense layer followed by a linear regression layer for output
	Weighted fusion of convolutional outputs from frames, audio data and transformers output of transcript with weights 0.1 for each frame, 0.4 for audio and 0.3 for transcript
XceptionResNet	Learning rate = 0.001
	Weighted fusion of convolutional outputs from frames, audio data and transformers output of transcript with weights 0.2 for each frame, 0.3 for audio and 0.1 for transcript

- *Root mean squared error (RMSE)* It is the square root of the mean squared error and is a better measure for the accuracy of a model.

These metrics collectively offer a comprehensive view of the performance of our multimodal models and enable us to draw meaningful conclusions regarding their effectiveness in personality trait prediction across various modalities.

5.1 Performance across modalities

We begin by analysing the performance of our models across various data modalities. Table 2 summarizes the results for different models and modalities:

From Table 2, we observe that the performance of the models varies across modalities. For text, DeBERTa performs poorly in terms of R2 score, while for audio and video, EfficientResNet and InceptionV3 perform well, respectively. When considering all modalities, XceptionResNet + BERT stands out with the highest accuracy, a high R2 score, and the lowest mean squared error and root mean squared error. It appears to be the best-performing model for the task involving all modalities. The bar chart in Fig. 9 illustrates the accuracy of various deep learning models in processing text, image, and frame data modalities for personality analysis, with EfficientResNet showing a leading performance.

5.2 Results on text modality

The results for two different models namely BERT and DeBERTa in terms of average accuracy and accuracy on various personality traits are in Table 3 given.

Both models demonstrate high average accuracy, with BERT scoring 88.09% and DeBERTa scoring 88.34%.

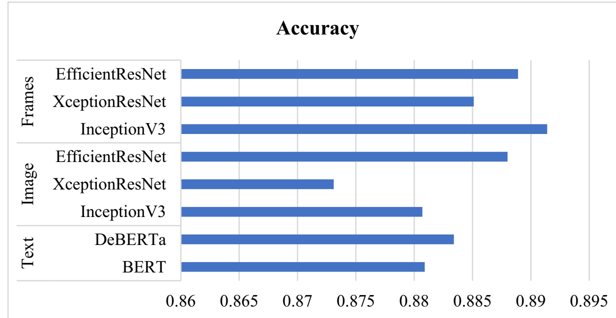

Fig. 9 Model comparison for personality analysis: accuracy of multimodal data processing

This suggests that both models perform well on the overall task of personality trait prediction from text data. Analysing the trait-wise accuracies, we find that Extraversion is the easiest trait to identify using the transcripts of the videos. This implies that individuals' spoken words can effectively predict whether they are extroverts or introverts. However, the other traits are relatively more challenging to identify from text data. Notably, Openness proves to be the most difficult to predict, likely requiring the assessment of body language and non-verbal cues.

5.3 Results on image modality generated from audio files

Moving on to the image modality derived from audio files, we summarize the results in Table 4:

In this modality, Extraversion remains the most predictable trait, closely followed by Openness. Conscientiousness appears to be the most challenging trait to identify. EfficientResNet appears to perform the best overall, with the highest average accuracy and the highest accuracy on most traits.

Table 2 Performance across modalities

Modalities	Model	Accuracy	R2 Score	Mean squared error	Root mean
Text	BERT	0.8809	− 2.60	0.0216	0.1469
	DeBERTa	0.8834	-8.31×10^9	0.0215	0.1467
Image	InceptionV3	0.8807	− 1.62%	0.0219	0.1478
	EfficientResNet	0.8880	9.97%	0.0193	0.1390
	XceptionResNet	0.8731	− 14.46%	0.0247	0.1569
Frames	InceptionV3	0.8914	15.37%	0.0182	0.1347
	EfficientResNet	0.8889	12.12%	0.0189	0.1374
	XceptionResNet	0.8851	4.64%	0.0205	0.1431
Fusion/ Decision models	InceptionV3 + BERT	0.8905	13.05%	0.0187	0.1366
	EfficientResNet + BERT	0.8876	10.55%	0.0194	0.1394
	XceptionResNet + BERT	**0.9212**	**54.49%**	**0.0098**	**0.0992**

The bold values indicate the best performing model

Table 3 Performance of text modality models

Model	Avg. Accuracy (%)	Accuracy on various traits				
		Extraversion (%)	Openness (%)	Agreeableness (%)	Neuroticism (%)	Conscientiousness (%)
BERT	88.09	89.24	87.25	87.85	87.81	88.28
DeBERTa	88.34	89.49	88.23	87.77	87.70	88.52

Table 4 Performance of audio-generated images models

Model	Avg. accuracy (%)	Accuracy on various traits				
		Extraversion	(%)Openness	(%)Agreeableness (%)	Neuroticism	(%)Conscientiousness (%)
InceptionV3	88.07	89.42	88.47	87.40	87.67	87.38
EfficientResNet	88.80	89.98	89.23	88.29	88.37	88.13
XceptionResNet	87.31	88.70	87.76	86.36	87.01	86.73

5.4 Results on frames generated from video files

For frames generated from video files, the results are summarized in Table 5:

In this modality, InceptionV3 delivers the best performance, achieving the highest average accuracy and excelling in most individual traits. Extraversion and Openness continue to be the easiest traits to identify from First Impressions, while traits like conscientiousness and neuroticism, which depend on one's approach to various life domains, remain challenging to judge from initial encounters.

5.5 Comparative analysis of the results

Inconsistent results are obtained when traits are analysed using multimodal decision models, as some of the models can analyse all the modalities together, while some of them fail to do so. This analysis (Table 6) shows that InceptionV3 + BERT Architecture is biased towards the video modality, hence showing results like those of Video models, while EfficientResNet + BERT architecture shows results closer to that of text models. The only model which could fit to all the modalities the best was XceptionResNet + BERT, which showed almost same accuracy on all labels except Neuroticism, thus making it the toughest trait to identify when we can fully assess the First Impression from a video, undertaking all its modalities.

Our study provides valuable insights into the effectiveness of multimodal models for predicting personality traits. The choice of data modality significantly impacts model performance, with multimodal fusion models showing superior accuracy. This suggests that combining textual, visual, and auditory information offers a holistic understanding of an individual's personality. The ease of predicting certain traits, such as Extraversion, from textual and visual data implies that linguistic and non-verbal cues play a pivotal role in personality assessment. On the other hand, traits like Openness, Conscientiousness, and Neuroticism may require a more comprehensive assessment, considering multiple modalities. The multimodal fusion architecture, XceptionResNet + BERT, stands out as a robust choice for personality trait prediction, with consistently high accuracy across all traits. However, careful consideration of the modality and task at hand is essential when choosing the most appropriate model architecture.

Table 5 Performance of video frames

Model	Avg. Accuracy (%)	Accuracy on various traits				
		Extraversion	(%)Openness	(%)Agreeableness (%)	Neuroticism	(%)Conscientiousness (%)
InceptionV3	89.14%	89.93	89.66	88.63	88.89	88.59
EfficientResNet	88.89%	89.99	89.28	88.35	88.61	88.21
XceptionResNet	88.51%	89.54	88.71	87.78	88.64	87.90

Table 6 Performance of multimodal decision models

Model	Avg. accuracy (%)	Accuracy on various traits				
		Extraversion	(%)Openness (%)	Agreeableness (%)	Neuroticism (%)	Conscientiousness (%)
InceptionV3 + BERT	89.05	89.94	89.56	88.46	88.78	88.49
EfficientResNet + BERT	88.76	89.56	89.09	88.65	88.08	88.39
XceptionResNet + BERT	92.12	92.35	92.24	92.16	91.31	92.51

Our system, leveraging the First Impression dataset, introduces several advancements in the field of personality-driven employment:

- *Enhanced accuracy in personality assessment* Utilizing multimodal data analysis, the system provides a more accurate and holistic view of an individual's personality traits, improving the precision of job-role matching.
- *Dynamic adaptability* By integrating real-time labour market analytics, the system can adapt to evolving job market trends, ensuring continuous alignment with current employment needs.
- *Scalable and comprehensive analysis* The architecture is designed to handle large datasets, allowing for a comprehensive analysis of personality traits across diverse job sectors.

5.6 Limitations and Ethical Concerns

While AI-driven personality analysis offers transformative potential for employment strategies, it is imperative to address the associated limitations and ethical concerns comprehensively:

- *Privacy and data protection* The use of AI for personality analysis raises significant privacy concerns, particularly regarding the collection, storage, and processing of sensitive personal data. Ensuring robust data protection measures and adhering to privacy regulations like GDPR are paramount to maintain individuals' trust and safeguard their personal information.
- *Fairness and bias* AI models can inadvertently perpetuate biases present in the training data, leading to unfair treatment of certain groups based on gender, race, ethnicity, or other characteristics. It is critical to implement measures to detect and mitigate such biases, ensuring that AI-driven personality assessments are fair and equitable across diverse populations.
- *Consent and transparency* Obtaining explicit consent from individuals whose data are used for personality analysis is essential. Furthermore, the processes and criteria used in AI-driven assessments should be transparent, allowing individuals to understand how their personality profiles are evaluated and used in employment decisions.
- *Regulatory compliance* Navigating the complex landscape of legal and ethical standards is a significant challenge. AI systems must be designed and operated in compliance with all applicable laws and ethical guidelines to avoid legal repercussions and maintain public confidence.
- *Moral and ethical responsibility* Beyond legal compliance, there is a moral obligation to use AI in a way that respects individual dignity and promotes positive outcomes for society. This includes careful consideration of how personality analysis can impact individuals' employment opportunities and career development.

In conclusion, while AI-driven personality analysis in employment contexts offers promising benefits, it is crucial to address these limitations and ethical concerns proactively. Doing so will not only enhance the fairness and effectiveness of these systems but also ensure their sustainability and acceptance in the long term.

6 Hyper-personalized employment in urban hubs: the role of personality traits across industries

The balance between traits and skills is crucial as shown in Fig. 10. Traits may provide the foundation for potential, while skills are often the practical tools through which one can execute tasks effectively. In a professional setting, both are important: traits can drive one's approach to work and interpersonal interactions, while skills are necessary to perform specific job functions. Together, they contribute to an individual's overall competency and potential for success in their career.

This section presents a series of case studies which underscore the universal importance of recognizing and nurturing the Big Five personality traits across various industries. In each sector, from technology to banking,

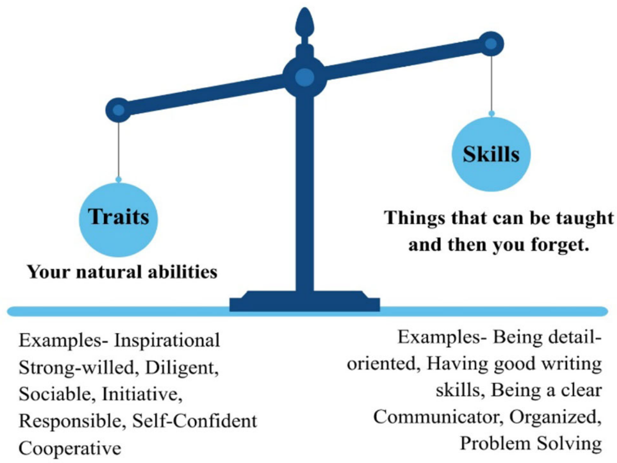

Fig.10 Future of urban employment: a balanced approach

healthcare to retail, the personalities of individuals significantly influence their effectiveness in their respective roles.

6.1 Case study 1: the role of personality traits in retail customer service and visual merchandising

In the dynamic landscape of the retail industry, success hinges not only on products and services but also on the human factor. Customer service and visual merchandising are two critical components that significantly impact a retail business's bottom line [22]. However, what often goes overlooked are the specific personality traits that make individuals excel in these roles. In this case study, we examine the relevance of the Big Five personality traits—Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness—in the contexts of Customer Service Representatives (CSRs) and Visual Merchandisers. The mapping of personality traits to retail roles is as follows:

- Extraversion

a) *CSR* Extraversion is valuable for CSRs. Enthusiastic and sociable CSRs engage customers warmly, creating a positive shopping experience.
b) *Visual merchandiser* While less critical, a touch of extraversion can benefit Visual Merchandisers in team collaborations and discussions about store aesthetics.

- Agreeableness

a) *CSR* Agreeableness is essential for CSRs. Friendliness, empathy, and a cooperative attitude are crucial for resolving customer issues and building rapport.
b) *Visual merchandiser* Agreeableness can help Visual Merchandisers collaborate effectively within the team and understand the preferences of diverse customers.

- Conscientiousness

a) *CSR* Attention to detail and reliability are vital for CSRs. They ensure accurate transaction processing and effective inventory management.
b) *Visual merchandiser* Conscientiousness plays a pivotal role in visual merchandising. Attention to detail ensures that displays are impeccable and on-brand.

- Neuroticism

a) *CSR* Low neuroticism is beneficial for CSRs. It helps them remain calm and composed, even in challenging situations.
b) *Visual Merchandiser* A moderate level of neuroticism can drive perfectionism in Visual Merchandisers, ensuring meticulous attention to visual details.

- Openness

a) *CSR* Openness is not as crucial for CSRs but can facilitate adaptability to various customer needs.
b) *Visual merchandiser* Openness to new design ideas and trends is vital for Visual Merchandisers to create innovative and visually captivating displays.

6.2 Case study 2: the role of personality traits in healthcare—nurses and hospital Administrators

The healthcare industry is a demanding and complex field that relies not only on technical skills but also on interpersonal abilities and emotional intelligence. In this case study, we examine how the Big Five personality traits—Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness—play a crucial role in the professions of Nurses and Hospital Administrators [23]. The mapping of personality traits to healthcare roles is as follows:

- Extraversion

a) *Nurse* Extraversion can benefit nurses in their interactions with patients and colleagues. A warm and sociable demeanour can help patients feel at ease and foster teamwork among healthcare staff.

b) *Hospital Administrator* While not as essential as for nurses, some level of extraversion can aid hospital administrators in effective communication with staff and patients, especially during crisis management.

- Agreeableness

a) *Nurse* Agreeableness is vital for nurses as it facilitates empathy and compassion when caring for patients. Patients often appreciate nurses who are kind, patient, and understanding.
b) *Hospital Administrator* Hospital administrators benefit from agreeableness when working with diverse healthcare teams and ensuring a harmonious working environment.

- Conscientiousness

a) *Nurse* Conscientiousness is a fundamental trait for nurses. Attention to detail, adherence to protocols, and organization are essential for providing high-quality patient care.
b) *Hospital Administrator* Hospital administrators rely heavily on conscientiousness to manage resources efficiently, maintain compliance with regulations, and uphold high standards of patient safety and care.

- Neuroticism

a) *Nurse* Lower levels of neuroticism can be advantageous for nurses as they often work in high-stress environments. Remaining calm and composed during critical situations is crucial.
b) *Hospital Administrator* Hospital administrators with lower neuroticism can handle the pressures of healthcare management effectively, ensuring the smooth operation of the hospital.

- Openness

a) *Nurse* Openness to new medical research and treatment methods can enhance a nurse's ability to provide the latest evidence-based care to patients.
b) *Hospital Administrator* Openness can be beneficial for hospital administrators when considering innovative healthcare practices, technologies, and management strategies.

6.3 Case study: the role of personality traits in banking—customer service representatives and risk analysts

The banking sector is a dynamic industry that relies not only on financial acumen but also on interpersonal skills and risk management [24]. This case study explores how the Big Five personality traits—Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness—play a pivotal role in the professions of Customer Service Representatives and Risk Analysts within the banking industry. The mapping of personality traits to banking roles is as follows:

- Extraversion

a) *Customer Service Representative* Extraversion can be a valuable trait for customer service representatives. An outgoing and approachable demeanour can enhance interactions with customers, leading to better customer satisfaction and loyalty.
b) *Risk Analyst* While not as crucial as for customer service representatives, some level of extraversion can aid risk analysts in effectively communicating their findings and insights to the bank's management or clients.

- Agreeableness

a) *Customer Service Representative* Agreeableness is vital for customer service representatives as it enables them to handle customer inquiries and complaints with empathy and patience, fostering positive customer relationships.
b) *Risk Analyst* Risk analysts benefit from agreeableness when working within a team to assess and mitigate financial risks, ensuring smooth collaboration.

- Conscientiousness

a) *Customer Service Representative* Conscientiousness is a fundamental trait for customer service representatives. Attention to detail, timeliness, and accuracy are essential when dealing with customers' financial transactions and inquiries.
b) *Risk Analyst* Conscientiousness is equally critical for risk analysts as it influences their ability to conduct thorough risk assessments, maintain meticulous records, and adhere to regulatory compliance.

- Neuroticism

a) *Customer Service Representative* Lower levels of neuroticism can be advantageous for customer service representatives as they often handle high-stress customer interactions. Remaining calm under pressure is vital.
b) *Risk Analyst* Risk analysts with lower neuroticism can objectively evaluate financial risks without being influenced by emotions or anxiety, ensuring more accurate risk assessments.

- Openness

a) *Customer Service Representative* While not as crucial as other traits, openness can benefit customer service representatives by enabling them to adapt to changing banking procedures and technologies.
b) *Risk Analyst* Openness can be beneficial for risk analysts when considering innovative risk assessment methodologies and staying updated with evolving financial markets.

6.4 Case study: the role of personality traits in the technology sector—software developers and UX designers

The technology sector is known for its innovation and dynamic nature. In this case study, we will examine how the Big Five personality traits—Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness—influence the work of Software Developers and User Experience (UX) Designers [25]. The mapping of personality traits to tech roles is as follows:

- Extraversion

a) *Software Developer* While software development often involves independent work and coding, some level of extraversion can be beneficial for effective collaboration with team members, especially in Agile environments.
b) *UX Designer* Extraversion can help UX designers when conducting user interviews and gathering feedback. Being able to engage with users and stakeholders can lead to better-designed products.

- Agreeableness

a) *Software Developer* Agreeableness can contribute to harmonious teamwork within development teams, fostering a positive and collaborative work environment.
b) *UX Designer* UX designers with high agreeableness can communicate more effectively with team members and stakeholders, facilitating the incorporation of feedback into the design process.

- Conscientiousness

a) *Software Developer* Conscientiousness is highly valuable for software developers. Attention to detail, commitment to deadlines, and thorough code reviews are essential for creating high-quality software.
b) *UX Designer* Conscientiousness ensures that UX designers meticulously plan and execute user research, usability testing, and design iterations, leading to user-friendly interfaces.

- Neuroticism

a) *Software Developer* Lower neuroticism can be advantageous for software developers as it helps them remain calm under pressure and adapt to rapidly changing technologies.
b) *UX Designer* UX designers with lower neuroticism can handle design critique and user feedback more constructively, ensuring the iterative design process remains productive.

- Openness

a) *Software Developer* Openness is valuable for software developers when exploring new programming languages, frameworks, and approaches to problem-solving.
b) *UX Designer* UX designers with openness are more likely to embrace innovative design concepts and adapt to emerging design trends, resulting in more creative and user-centric solutions.

These case study highlights the importance of recognizing and nurturing the Big Five personality traits that serve as invaluable indicators for role alignment and workforce optimization, empowering organizations across industries to achieve success and foster innovation.

6.5 Economic impact analysis

The deployment of our AI-driven hyper-personalized job matching system is anticipated to have substantial economic implications for the job market and organizational productivity.

- *Job Market Efficiency* By aligning candidates to roles that resonate with their innate traits and acquired skills, the system can significantly streamline the hiring process, reducing the time and resources spent on finding the right candidate for the job.
- *Employee Turnover Rates* With a more precise match between job roles and personal traits, employee satisfaction is likely to increase, leading to a reduction in turnover rates. This can result in considerable cost savings for organizations, as the expenses associated with recruiting and training new employees are lowered.
- *Overall Productivity* Employees who are well-matched to their roles are more likely to be engaged and motivated, enhancing productivity. This improved efficiency can contribute to the organization's bottom line and economic growth in the broader market.

6.6 Technical implementation: addressing scalability and logistical challenges

The technical execution of our hyper-personalized job matching system, while feasible, necessitates a detailed examination of scalability and logistical considerations to ensure successful real-world application:

- *Scalability Concerns* Technically, the system must be designed to efficiently handle varying scales of operation, from small enterprises to large multinational corporations. This involves optimizing the system architecture for load handling, data processing speed, and storage capacity, ensuring that it can expand or contract based on the organization's size and requirements.
- *Logistical Integration* From a logistical standpoint, integrating the system into existing human resources and IT infrastructures is paramount. This includes establishing seamless data pipelines, ensuring compatibility with current HR software, and providing mechanisms for continuous data update and retrieval.
- *Resource Allocation* Adequate allocation of computational and human resources is essential for maintaining system performance. This involves planning for the necessary hardware, software, and personnel to manage the system effectively and respond to any issues promptly.
- *System Maintenance and Support* Ongoing technical support and maintenance are crucial for long-term system stability and reliability. Regular updates, bug fixes, and system audits should be planned to ensure the system remains current and effective.
- *Training and Adaptation* To facilitate logistical adoption, comprehensive training programs for HR professionals and end-users are necessary. This will aid in smooth system integration and ensure that users are proficient in leveraging the system's capabilities.

Addressing these technical and logistical elements will not only enhance the system's practical deployment but also ensure its sustainable operation and scalability in diverse organizational environments.

7 Conclusion

In the context of smart cities, this research addresses the shift in employment strategies towards greater personalization, with the goal of enhancing job satisfaction and economic efficiency. The study utilizes multimodal deep learning techniques to create a hyper-personalized job matching system based on individual personality traits, drawing on the First Impressions V2 dataset. Among various model architectures, the fusion of XceptionResNet with BERT demonstrates exceptional performance, achieving a remarkable accuracy of 92.12% and other robust metrics. These empirical findings underscore the efficacy of the XceptionResNet + BERT approach in personality trait mapping, presenting an innovative and efficient method for job matching in urban environments. This research has the potential to revolutionize recruitment practices in smart cities by ensuring job placements that align not only with skills but also with individual personality traits, thereby optimizing both individual job satisfaction and organizational productivity. Theoretical case studies across diverse sectors further illustrate the model's potential to enhance both individual and organizational outcomes in smart city settings.

The limitations of this research lie in its dataset specificity, ethical considerations, interpretability, and potential biases, which should be mitigated in future studies. Future work should encompass real-world validation, multimodal enhancement, dynamic personality assessment, bias mitigation strategies, user-centred design, cross-cultural adaptation, and regulatory compliance to refine and extend the model's applicability. By addressing these limitations and pursuing these avenues, the research can contribute to the development of a more robust, fair, and globally adaptable personality-based job matching system, thus advancing the evolution of modern employment strategies within the dynamic landscape of smart cities and the Fourth Industrial Revolution.

Funding No funding was received for this research.

Data availability A publicly available dataset is used in this research; dataset will be made available on request.

Declarations

Conflict of Interest The authors have no competing interests related to the work submitted for publication.

References

1. Rose, G. (Ed.). (2016). *The fourth industrial revolution: A Davos reader*. Foreign Affairs.
2. Purcell KPJ (1998) In-sourcing, outsourcing, and the growth of contingent labour as evidence of flexible employment strategies. Eur J Work Organ Psy 7(1):39–59
3. Ponce-López V, Chen B, Oliu M, Corneanu C, Clapés A, Guyon I, Baró X, Escalante HJ, Escalera S. (2016). Chalearn lap 2016: First round challenge on first impressions-dataset and results. In: *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8–10 and 15–16, 2016, Proceedings, Part III 14* (pp. 400–418). Springer International Publishing.
4. Nirthika R, Manivannan S, Ramanan A, Wang R (2022) Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study. Neural Comput Appl 34(7):5321–5347
5. Wang J, Li H, Wang L, Wu C (2023) A multimodal dialogue system for improving user satisfaction via knowledge-enriched response and image recommendation. Neural Comput Appl 35(18):13187–13206
6. Pianesi F, Mana N, Cappelletti A, Lepri B, Zancanaro M (2008). Multimodal recognition of personality traits in social interactions. In: Proceedings of the 10th international conference on Multimodal interfaces (pp. 53–60).
7. Alam F and Riccardi G. (2014). Predicting personality traits using multimodal information. In Proceedings of the 2014 ACM multi media on workshop on computational personality recognition (pp. 15–18).
8. Escalante HJ, Kaya H, Salah AA, Escalera S, Güçlütürk Y, Güçlü U, Baró X, Guyon I, Junior JC, Madadi M, Ayache S (2018). Explaining first impressions: Modeling, recognizing, and explaining apparent personality from videos. *arXiv preprint* arXiv:1802.00745.
9. Gorbova J, Avots E, Lüsi I, Fishel M, Escalera S, Anbarjafari G (2018) Integrating vision and language for first-impression personality analysis. IEEE Multimedia 25(2):24–33
10 Junior JCJ, Güçlütürk Y, Pérez M, Güçlü U, Andujar C, Baró X, Escalera S (2019) First impressions: A survey on vision-based apparent personality trait analysis. IEEE Trans Affect Comput 13(1):75–95
11. Wingate TG, Bourdage JS (2019) Liar at first sight? Early impressions and interviewer judgments, attributions, and false perceptions of faking. J Pers Psychol 18(4):177
12. Nørskov S, Damholdt MF, Ulhøi JP, Jensen MB, Mathiasen MK, Ess CM, Seibt J (2022) Employers' and applicants' fairness perceptions in job interviews: using a teleoperated robot as a fair proxy. Technol Forecast Soc Chang 179:121641
13. Aslan S, Güdükbay U, Dibeklioğlu H (2021) Multimodal assessment of apparent personality using feature attention and error consistency constraint. Image Vis Comput 110:104163
14 Song S, Jaiswal S, Sanchez E, Tzimiropoulos G, Shen L, Valstar M (2021) Self-supervised learning of person-specific facial dynamics for automatic personality recognition. IEEE Trans Affect Comput. 14(1):178–195
15. Demir A, Yilmaz F, Kose O (2019). Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3. In: *2019 medical technologies congress (TIPTEKNO)* (pp. 1–4). IEEE.
16. Canayaz M (2021) C+ EffxNet: a novel hybrid approach for COVID-19 diagnosis on CT images based on CBAM and EfficientNet. Chaos, Solitons Fractals 151:111310
17. Panthakkan A, Anzar SM, Jamal S, Mansoor W (2022) Concatenated Xception-ResNet50—A novel hybrid approach for accurate skin cancer prediction. Comput Biol Med 150:106170
18. Jain D, Kumar A, Beniwal R (2022, July). Personality BERT: A Transformer-Based Model for Personality Detection from Textual Data. In: *Proceedings of International Conference on Computing and Communication Networks: ICCCN 2021* (pp. 515–522). Singapore: Springer Nature Singapore.
19. He P, Liu X, Gao J, Chen W. (2020). Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint* arXiv:2006.03654.
20 Jain D, Kumar A (2024) AI unveiled personalities: profiling optimistic and pessimistic attitudes in Hindi dataset using transformer-based models. Expert Syst. https://doi.org/10.1111/exsy.13572
21 Handelman GS, Kok HK, Chandra RV, Razavi AH, Huang S, Brooks M, Asadi H (2019) Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. Am J Roentgenol 212(1):38–43
22. Huang JL, Ryan AM (2011) Beyond personality traits: a study of personality states and situational contingencies in customer service jobs. Pers Psychol 64(2):451–488
23. Ganu D, Kogutu CN (2014) Effect of the big five personality traits on job satisfaction and organizational commitment in the healthcare industry: the case of Kenya. Am J Health Sci (AJHS) 5(2):145–154
24. Rashid NM, Sah NFM, Ariffin NM, Ghani WSWA, Yunus NSNM (2016) The Influence of Bank's frontlines' personality traits on job performance. Procedia Econ Financ 37:65–72
25. Yilmaz M, O'Connor RV, Colomo-Palacios R, Clarke P (2017) An examination of personality traits and how they impact on software development teams. Inf Softw Technol 86:101–122

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.