Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Architecture for IT Internship Recruitment Process Based on AWS Cloud

Rotimi-Williams Bello, Pius A. Owolawi, Etienne A. van Wyk, Chunling Tu
Department of Computer Systems Engineering, Faculty of Information and Communication Technology, Tshwane University of Technology, South Africa; email: bellorw@tut.ac.za

ARTICLE INFO

Received: 18 Dec 2024

Revised: 10 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

Introduction: The current recruitment processes face challenges due to lack of automation and scalability. Filtering student data manually is both time-intensive and prone to errors. Managing high volumes of applications securely is difficult, and there is limited capability to effectively track applications and notify students.

Objectives: This paper aims to design a secure, automated, and scalable cloud application on Amazon Web Services (AWS) leveraged by NLP-based deep learning models like Bidirectional Encoder Representations from Transformers (BERT) and Sentence Transformer to streamline Information Technology (IT) internship recruitment using Amazon S3 + CloudFront, AWS WAF, API Gateway, AWS Step Functions, AWS Lambda, Amazon RDS, AWS SNS, and CloudWatch.

Methods: AWS's functions as well as powerful machine learning tools were utilized for the recruitment process automation and streamlining. We present BERT fine-tuning results on the research tasks by employing datasets such as (1) Stanford Question Answering Dataset (SQuAD v1.1), (2) SQuAD 2.0, (3) General Language Understanding Evaluation (GLUE) benchmark, and (4) Situations With Adversarial Generations (SWAG) dataset.

Results: With SQuAD v1.1, our proposed AWS-based method obtained 88.1 (EM) and 95.1 (F1) for Dev, and 90.1 (EM) and 95.2 (F1) for Test. With SQuAD v2.0, our proposed AWS-based method obtained 83.2 (EM) and 86.3 (F1) for Dev, and 83.4 (EM) and 88.4 (F1) for Test. Our proposed AWS-based method obtained 88.9 (Dev) and 88.8 (Test) for SWAG Dev and test accuracy.

Conclusions: The solution provided by our proposed method simplifies the recruitment process, strengthens security through AWS services, scales effortlessly to manage high volumes of applications, automates notifications for better communication, provides administrators with convenient access to recruitment data, offers a cost-efficient and fully managed cloud-based infrastructure.

Keywords: Amazon Web Services, Automated System, Information Technology, Internship Recruitment.

1. INTRODUCTION

Considerable growth has been witnessed in the information technology (IT) sector in using services, software, hardware, and infrastructure driven by clouds, especially those services that are not within a company's own facilities [1]. Flexible architecture has been the driving force of many institutions, enabling institutions' scaling up or scaling down as needed, leveraging utilities like computation and storage, managing and recovery from disaster in the event of service downtime. Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform are the three giant companies in the cloud business. However, AWS is in the forefront when it has to do with market share [2]. AWS is free to access and has a robust platform. AWS services are considered the highest ranked due to their capability for execution and completeness of vision with their patented Magic Quadrant analysis [3]. A study conducted by some IT professionals reveals a high demand in cloud skills in addition to the rising stature of AWS.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

The results are an indication of AWS acceptance by the populace and a shift away from Google compared to prior results [4]. AWS provides a suite of tools for launching server and serverless resources, file and blob storage systems, and fully managed databases, in addition to emerging technologies like artificial intelligence (AI), quantum computing, and the Internet of Things (IoT) [5]. AWS, as one of the leading providers of cloud services, offers various tools and services such as EC2 (for hosting web applications), S3 (for storing resumes), RDS (for managing databases), Lambda (supports building scalable APIs), API Gateway (for integrating with other recruitment tools like LinkedIn, etc.), and AWS Shield and WAF (Web Application Firewall) (for security against Distributed Denial-of-Service (DDoS) attacks or malicious access), that can be leveraged to build efficient recruitment platforms. Through the AWS cloud, organizations can accommodate scalable applications that are readily available with low latency. Companies or organizations are continuously seeking to hire students or graduates for IT internship positions through IT internship recruitment scheme [6].

Natural Language Processing (NLP) as a field of AI focuses on the association between humans and computers through natural language. Human language can be understood, interpreted, and generated by computers using NLP in a meaningful and useful way. Various tasks such as resume parsing, candidate screening, job description analysis, sentiment and emotional tone analysis, interview transcription and analysis, chatbots for candidate interaction, and candidate evaluation from open-ended responses, can be automated and improved by NLP in the context of the process for IT internship recruitment based on AWS Cloud or any related recruitment systems. The techniques that are commonly used for NLP implementation in recruitment and internship processes are: tokenization, text classification, named entity recognition (NER), part-of-speech tagging (POS), sentiment analysis, topic modeling, dependency parsing, word embeddings, and text summarization. Bidirectional Encoder Representations from Transformers (BERT) and Sentence Transformer are NLP-based deep learning models designed for human language processing and analysis.

Although most of these internships are typically short in duration, their design provides immeasurable practical work experience, skills development, and exposure to the IT industry for individuals seeking a career in IT [7]. The steps in recruitment usually include job posting and advertisement, application, screening and interviews, offer and onboarding, and internship [8]. Both the company and the intern benefit from internships in IT. Companies acquire new perspectives and talent, while interns acquire skills, network, and often have the transition opportunity into permanent roles post-internship. However, the current recruitment processes face challenges due to a lack of automation and scalability. Filtering student data manually is both time-intensive and prone to errors. Managing high volumes of applications securely is difficult, and there is limited capability to effectively track applications and notify students.

This paper aims to design a secure, automated, and scalable cloud application on Amazon Web Services (AWS) leveraged by NLP-based deep learning models like BERT and Sentence Transformer to streamline IT internship recruitment using Amazon S3 + CloudFront, AWS WAF, API Gateway, AWS Step Functions, AWS Lambda, Amazon RDS, AWS SNS, and CloudWatch. The goals are to implement automated filtering of student applications based on predefined criteria, enable seamless notification dispatch to students, ensure secure storage of accepted applications, and offer administrators straightforward access to recruitment outcomes. The solution provided by our proposed method simplifies the recruitment process, strengthens security through AWS services, scales effortlessly to manage high volumes of applications, automates notifications for better communication, provides administrators with convenient access to recruitment data, offers a cost-efficient and fully managed cloud-based infrastructure. Figure 1 shows an overview of architecture for IT internship recruitment process based on AWS cloud.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 1. Overview of architecture for IT internship recruitment process based on AWS cloud.

2. RELATED WORK

When considering architecture for IT internship recruitment process based on AWS cloud, there is a need for a system or platform capable of leveraging AWS cloud infrastructure for the streamlining, scaling, and enhancement of the recruitment process for IT internships. Among the areas such architecture may focus on are automation, data management and security, scalability and performance, development (Dev) and operations (Ops) for continuous integration and continuous deployment, and collaboration and tracking, while leveraging AWS services to improve overall experience. In this section, we review related work on IT internship recruitment process based on AWS cloud.

2.1. Cloud-based Recruitment Systems

Cloud computing has been adopted by many recruitment platforms to streamline processes such as management of candidates, scheduling of interview, and processing of application. Kumar [9] studied the cloud technology invasion and the benefits it offered in talent acquisition. The researcher identified and analyzed the recent technological changes and challenges in the health industry in the talent acquisitions process. Research has revealed that cloud-based solutions provide flexibility, scalability, and minimized IT overhead for recruitment organizations. Dong and Salwana [10] in order to improve financial, marketing, and collaborative performances of multinational organizations, proposed cloud-based human resource management (CBHRM) and cloud-based supply chain management (CBSCM). Structural equation modeling (SEM) was utilized in their study for estimation of validity and reliability of the measurement model, and assessment of the causal model. Statistical package for the social sciences (SPSS) and linear interrelationships structural relations (LISREL) were employed for analysis of the offered model and the questionnaires. The results they obtained showed the influence of CBSCM and CBHRM on the performance of the companies. The results validate their proposition that CBSCM and CBHRM are essential and necessary for financial, marketing, and collaborative performances.

2.2. AWS Services for Automation in Recruitment

The process of sending notifications to candidates could be automated using AWS Simple Notification Service (SNS). Moreover, interviews can be scheduled and recruiters notified about new applicants using the AWS SNS system. Email and message platforms can be integrated with this system. Ahmad et al. [11] proposed a machine learning solution based on NLP for automated scheduling and communication of interview in natural language. They developed a scalable and cost-effective solution by using two deep-learning NLP models as the system's foundation, while AWS serverless stack and Amazon SageMaker were employed for training and inference. Their work demonstrates NLP's potential for developing a smart solution for an everyday business problem.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Recruiters can use serverless architecture for execution of functions such as updating the status of candidates, sending interview invitations to applicants, and triggering responses automatically with no need for a dedicated server. Pandit et al. [12] introduced a framework based on web and designed for recruitment process revolutionization within the IT industry. Key recruitment functions were automated and enhanced by the framework using AI technologies such as NLP, machine learning, and data analytics. Core features of the framework include resume screening, which uses a ranking system and advanced algorithm for job description matching, whereby candidates can be efficiently shortlisted based on different parameters. Moreover, a Google Dialogflow-powered chatbot-based assessment system was incorporated by the framework to conduct skill-based evaluations.

2.3. Data Management and Security in Recruitment

Large amounts of structured and unstructured data are generated through recruitment processes. Relational Database Service (RDS) of AWS or NoSQL database service DynamoDB can be employed to store information about individual candidates, schedules of interview, job descriptions, and hiring progress, all with security, scalability, and reliability. Tobroni and Hamid [13] proposed a security-driven system that could help the human resources (HR) department in managing recruitment processes with capabilities for data retention. They incorporated automated filtering techniques based on job requirements, enabling the HR department to effectively manage job applications and identify potential job candidates. Moreover, the design of the system does not allow storage of personal data in the database beyond six months, complying with data retention guidelines.

Sensitive data includes personal data like resumes and interviews, and their handling requires reliable security, which is a major concern. Identity and Access Management (IAM) is provided by AWS for access control management, enabling recruiters to manage and control access to which resources, and ensuring compliance with regulations such as General Data Protection Regulation (GDPR) or Health Insurance Portability and Accountability Act (HIPAA). Roy et al. [14] formalized recruitment problem as the optimal recruitment problem (ORP), in which the goal is about selecting from a set of candidates, the minimum number of fresh employees, to occupy the available positions of the outgoing employees, while ensuring that the specified security conditions are sated.

For specification of authorization policies and constraints, they used Attribute-Based Access Control (ABAC) model because it is considered as de facto next-generation framework for managing security policies of organizations. Their research showed ORP problems as NP-hard; they proposed a solution based on greedy heuristic, wherein an extensive experimental evaluation reveals the performance of the proposed solution. Jeong and Choi [15] designed a platform for recruitment management using digital certificate on blockchain. This is also in line with the research conducted in [16] on managing and securing recruitment data.

2.4. Scalability and Performance in the Recruitment Process

Sudden traffic surges may affect recruitment systems, especially in peak hiring period. An automated Elastic Load Balancer (ELB) is a common technique for distributing traffic of incoming applications to multiple instances, ensuring performance consistency even during periods with high demand. Kurek et al. [17] carried out a study that explored Zero-Shot Recommendation AI Models application for matching process enhancement. They assessed the effectiveness of advanced pretrained models like all-MiniLM-L6-v2 in aligning job descriptions with candidate's profiles. They also applied similar metrics like cosine and dot product similarity for the same purpose. Their Top K Accuracy-based evaluation across various rankings showed a remarkable enhancement in matching accuracy better than the conventional methods.

Alagha et al. [18] proposed a novel Stable Data-based Recruitment System (SDRS) for localization tasks. The proposed model (1) incorporates a new recruitment parameter based on data for dynamic exploit of data readings that could guide the recruitment system into informative workers selection, while taking their mobility into consideration; (2) introduces a stable method for coverage assessment that takes range-free sensors and workers' mobility into consideration; and (3) incorporates an optimized approach for two-phase recruitment using greedy and genetic methods. The results obtained from their experiment demonstrate the efficiency and reliability of the proposed approach leading to a fast and quality localization outcome.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3. MATERIALS AND METHODS

We employ a method that belongs to NLP, a subfield of AI that enables the interaction between humans and computers using natural language. BERT and Sentence Transformer, NLP-based deep learning models designed for human language processing and analysis, were employed in this study. BERT, a widely recognized model for evaluating candidate responses, has been pre-trained and fine-tuned with a large amount of data and training data, respectively, making it suitable for the problems addressed in this study. The procedure for pre-training BERT largely followed the existing procedure on pre-training of language models. For the pre-training corpus, BooksCorpus (800M words) [19] and English Wikipedia (2,500M words) were used. For Wikipedia, only the text passages were extracted while the lists, tables, and headers were ignored.

In order to extract long contiguous sequences, a document-level corpus is preferred to a shuffled sentence-level corpus such as the Billion Word Benchmark [20]. The retraining and fine-tuning process of the proposed model were carried out using TensorFlow with Amazon Sage Maker. Figure 2 illustrates the architecture for IT internship recruitment process based on AWS cloud. To further boost the performance of the proposed models, this study integrated a second model, namely Sentence Transformer, which has capability for converting sentence into embedding, accelerating comparisons and analyses.

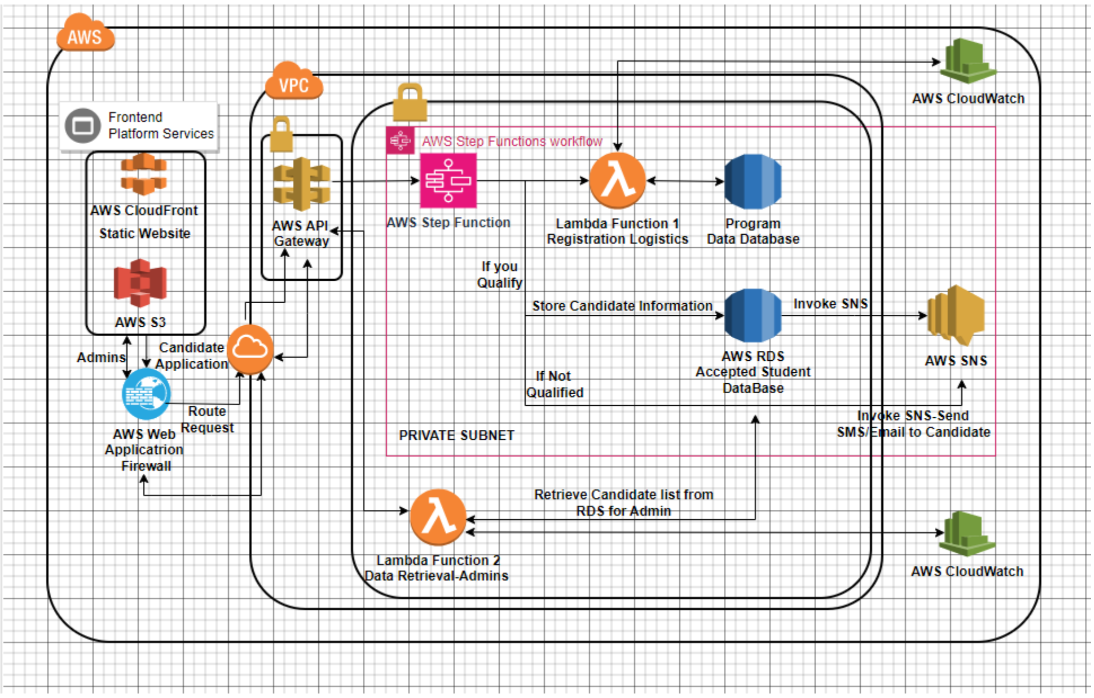

Figure 2. Architecture for IT internship recruitment process based on AWS cloud.

To integrate NLP into the AWS-based IT internship recruitment system proposed in this paper, we utilized several AWS services and functions including (a) Lambda Functions, for invoking the model for real-time inference and tasks automation such as parsing resumption or sending notifications based on candidate responses; (b) Amazon Comprehend, for insights analysis and extraction from candidate documents like resumes and interview transcripts; (c) Amazon Textract, for text and structured data extraction from scanned resumes or other documents; (d) Amazon Lex, for building chatbots for activity such as candidate interactions, interview scheduling, or frequently asked question (FAQ) handling; (e) Amazon SageMaker, for training and hosting BERT models, and training custom models and NLP model deployment specific to the recruitment needs; (f) Amazon Transcribe, for converting the content of spoken interview into text for further processing and analysis; (g) Amazon S3, for storing large amounts of resume data, job descriptions, and interview data; (h) Amazon DynamoDB, for storing and managing metadata, such as candidate information, interview schedules, and match scores; (i) Amazon QuickSight, for creating dashboards for HR personnel to review candidate rankings and match scores; (j) Amazon S3 + CloudFront, for hosting the frontend and ensuring global content delivery; (k) AWS WAF, for protecting the application from web-

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

based threats; (l) API Gateway, for managing APIs and routing incoming requests; (m) AWS Step Functions, for coordinating workflows seamlessly; (n) Amazon RDS, for storing program criteria and records of accepted students; (o) AWS SNS, for delivering acceptance or rejection notifications for applications. (p) CloudWatch, for providing monitoring and auditing capabilities.

Figure 3 shows the system flow of IT internship recruitment process based on AWS services, starting from the first stage when the students submit their details through a frontend hosted on Amazon S3 to the last stage when the administrators access results through a secure API powered by Lambda.

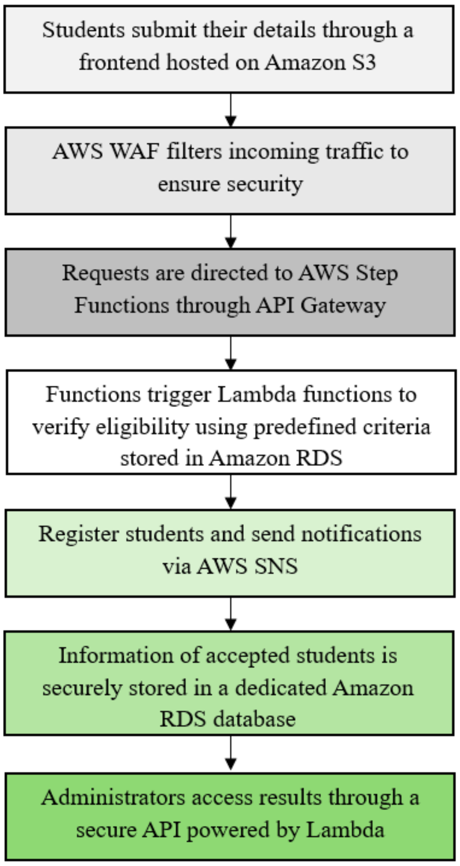

Figure 3. System flow of IT internship recruitment process based on AWS services.

3.1. BERT-based Approach

We implement the BERT-based approach on AWS Cloud by collecting data such as resumes in structured or unstructured formats like PDF and DOCX, etc., and we also gathered interviews that were transcribed or in text format, including the job descriptions of the IT internship roles, which were stored in a structured format like JSON. For data preprocessing, we converted unstructured data into structured data and tokenized the text in resumes and job descriptions. We fine-tuned a pretrained BERT model on labeled data to match resumes with job descriptions and computed similarity scores to rank candidates based on how well skills and resumes align with the job requirements. We deployed the BERT model on AWS SageMaker for model fine-tuning and service at scale. We uploaded the processed dataset to SageMaker and fine-tuned BERT using Hugging Face's transformers library framework.

We used AWS Lambda to trigger real-time processing by this model (e.g., when a new resume is uploaded to S3). For the sentiment analysis of the interview data (either text-based transcripts or voice-to-text), we used the BERT model to gauge the candidate's emotional tone and appropriateness in responses. In this study, a specific machine learning model was employed in evaluating the responses of candidates for each topic. Evaluation of candidates' responses

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

can be enhanced by implementing separate models for each issue. To achieve this, the questions and responses are transferred to a feedforward network for training, which rates the questions and responses accordingly. A training strategy was employed for the model (Figure 4). The use case of BERT-based approach is as follows:

Step 1: Resumes are uploaded to a web portal by a candidate.

Step 2: AWS Lambda invokes a SageMaker endpoint, which uses a fine-tuned BERT model to analyze the resume and compare it against the job description.

Step 3: A match score is assigned by the model and stored in DynamoDB before it is sent to the HR personnel for review.

Step 4: Sentiment analysis is employed in processing the interview transcripts if an interview is conducted, and the results are fed back into the system.

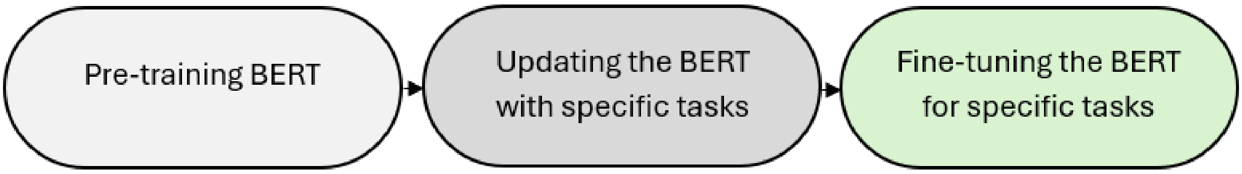

Figure 4. Bert based approach.

3.2. Sentence Transformer-based Approach

Pre-trained models like BERT can be leveraged by a Sentence Transformer-based approach for enhanced IT recruitment process on AWS Cloud. The need to automate and optimize candidate screening was the reason for using Sentence Transformers for the IT internship recruitment process. Sentence Transformer-based approach facilitates ranking and evaluation of suitably qualified applicants for a particular internship by comparing the content of their resumes, job descriptions, and interview answers. The BERT model generates contextualized word embeddings for each token in the input text when employed for question answering or text classification tasks. Pooling operations are often applied to the max pooling or mean pooling embeddings in order to obtain a fixed-size representation of the entire input.

For a student's answer evaluation, contextualized embeddings could be produced by the BERT model for each token in both the gold answer and the student's answer, and each answer’s embeddings could then be pooled for u and v production, representing the fixed-size vector of the gold answer and the student's answer, respectively. Figure 5 shows the sentence Transformer-based approach.

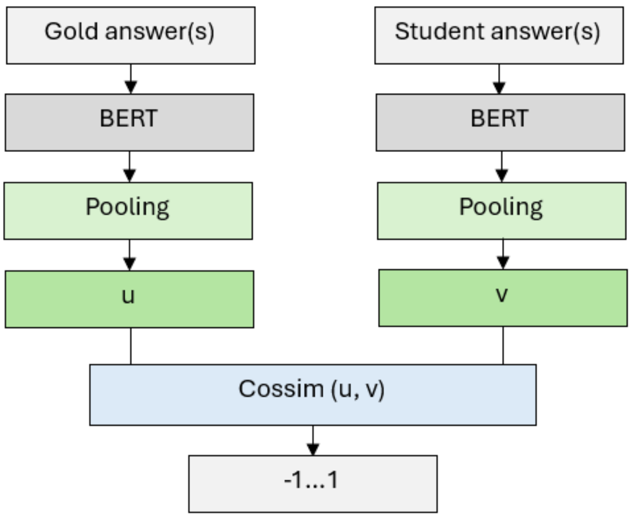

Figure 5. Sentence Transformer-based approach.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The similarity between the two answers is measured by the cosine similarity (Cossim) between their vectors and computed as follows.

$$\text{Cossim}\ (u, v) = (u \cdot v) / (||u||\ ||v||) \quad (1)$$

Where Cossim (u,v) denotes the cosine similarity between vectors u and v.

Cosine similarity, often used in natural language processing and information retrieval, measures the similarity between two non-zero vectors (also for comparing similarity between two documents or sets of words) by calculating the cosine of the angle between them, ranging from -1 to 1, where 1 indicates perfect similarity and -1 indicates perfect dissimilarity. Besides the BERT-based model, we utilized an additional model by which addition and assessment of new questions can be made by merely providing a reference answer, which the candidate's responses are compared to during the assessment process to determine if a score should be assigned based on the degree of similarity between the two.

This approach addresses the limitation of time and resources needed for training data creation by utilizing the pre-trained BERT model for response encoding and comparison using cosine similarity. The key components in this approach are AWS Cloud and Sentence Transformer models. AWS Cloud is the infrastructure responsible for scalability and deployment, and the functions include AWS Lambda for serverless execution, Amazon SageMaker for training models, AWS S3 for storing candidate resumes and interview transcripts, Amazon ElasticSearch or RDS for efficient storage and querying of candidate data, and AWS Comprehend for natural language processing tasks. Sentence Transformer models convert texts like resumes, job descriptions, and interview answers into embeddings numerical representations for semantic meaning capturing, allowing matching based on similarity.

4. EXPERIMENT AND RESULTS

The entire system is built on AWS serverless services using AWS Lambda for serverless inference, Amazon Dynamo DB for storing the embeddings, AWS CloudWatch for logging the recruitment process performance, monitoring API calls and tracking any errors, and Amazon SQS, with deep learning training and inference handled by Amazon Sage Maker.

We present BERT fine-tuning results on the research tasks by employing datasets such as (1) The Stanford Question Answering Dataset (SQuAD v1.1) comprising 100k crowdsourced question/answer pairs [21] and the model was fine-tuned on it for 3 epochs with a learning rate of 5e-5 and a batch size of 32; (2) The SQuAD 2.0 task, on which the model was fine-tuned for 2 epochs with a learning rate of 5e-5 and a batch size of 48, extends the SQuAD 1.1 problem definition with restriction of short answer in the provided paragraph, making the problem more realistic; (3) The General Language Understanding Evaluation (GLUE) benchmark [22], on which the model was fine-tuned for 3 epochs with best learning rate from among 5e-5, 4e-5, 3e-5, and 2e-5, and a batch size of 32 for all GLUE tasks, comprises a collection of diverse natural language understanding tasks including the following datasets: (a) Multi-Genre Natural Language Inference (MNLI) [23], (b) Quora Question Pairs (QQP) [24], (c) Question Natural Language Inference (QNLI), a version of SQuAD that has been transformed to a binary classification task [22], (d) The Stanford Sentiment Treebank (SST-2) [25], (e) The Corpus of Linguistic Acceptability (CoLA) [26], (f) The Semantic Textual Similarity Benchmark (STS-B) [27], (g) Microsoft Research Paraphrase Corpus (MRPC) [28], (h) Recognizing Textual Entailment (RTE) [29], and (i) Winograd NLI (WNLI) [30]; (4) The Situations With Adversarial Generations (SWAG) dataset on which the model was fine-tuned for 3 epochs with a learning rate of 2e-5 and a batch size of 16, comprises 113k sentence-pair completion examples that evaluate grounded commonsense inference [31].

Compared to pre-training, fine-tuning is relatively inexpensive. The fine-tuning results of Bert-based model, Sentence Transformer-based model, and AWS-based model are presented in Tables 1 to 4.

Table 1: SQuAD v1.1 Results

System	Dev		Test	
	EM	F1	EM	F1
BERT-based	87.3	93.3	88.6	94.3

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Sentence Transformer-based	87.9	94.2	89.1	94.8
AWS-based	88.1	95.1	90.1	95.2

Table 2: SQuAD v2.0 Results

System	**Dev**		**Test**	
	EM	F1	EM	F1
BERT-based	81.3	84.1	81.6	87.3
Sentence Transformer-based	82.1	85.4	82.2	87.7
AWS-based	83.2	86.3	83.4	88.4

Table 3: GLUE Test Results

System	**MNLI (m/mm)**	**QQP**	**QNLI**	**SST-2**	**CoLA**	**STS-B**	**MRPC**	**RTE**
BERT-based	87.6/86.5	73.2	93.6	95.1	61.7	87.6	90.7	71.1
Sentence Transformer-based	88.6/87.5	74.1	94.7	96.3	63.1	88.8	91.3	71.8
AWS-based	89.6/88.5	75.1	95.6	96.2	62.9	88.9	91.5	72.3

Table 4: SWAG Dev and Test Accuracy Results

System	**Dev**	**Test**
BERT-based	87.6	87.1
Sentence Transformer-based	88.5	88.3
AWS-based	88.9	88.8

In this study, we designed a secure, automated, and scalable cloud application on Amazon Web Services (AWS) leveraged by NLP-based deep learning models like BERT and Sentence Transformer to streamline IT internship recruitment. A fine-tuned BERT-based model was employed in making the first model, while a transformer-based model was employed in making the second model. Both fine-tuned models were employed in training the proposed AWS model for questioning and evaluation of candidate responses to technical questions. The results presented in Tables 1 to 4 showed that the proposed AWS model was effective for all the tasks and evaluation of candidate responses to technical questions. AWS serverless stack and training/inference were employed in building the system using Amazon SageMaker, thereby allowing for quick scaling at minimal cost.

CONCLUSIONS

The AWS cloud architecture for the IT internship recruitment program provides a secure, scalable, and automated solution to overcome recruitment challenges. Core services like API Gateway, Step Functions, and Lambda ensure streamlined application processing, instant notifications, and efficient candidate management. Security is reinforced through AWS WAF and VPC, while CloudWatch delivers robust monitoring and auditing capabilities. By utilizing AWS's fully managed services, the solution minimizes operational burdens, enhances efficiency, and ensures

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

seamless experience for students and administrators alike. This system demonstrates the potential of cloud computing to revolutionize recruitment processes, laying the groundwork for future scalability and innovation. BERT is a powerful model, understanding how it arrives at decisions can be challenging. In future, we hope to diversify and balance the dataset used in training the BERT model for better understanding, predictions, and results.

REFERENCES

[1] Sunyaev, A., & Sunyaev, A. (2020). Cloud computing. Internet computing: Principles of distributed systems and emerging internet-based technologies, 195-236.

[2] Gupta, B., Mittal, P., & Mufti, T. (2021). A review on amazon web service (aws), microsoft azure & google cloud platform (gcp) services. In Proceedings of the 2nd International Conference on ICT for Digital, Smart, and Sustainable Development, ICIDSSD 2020, 27-28 February 2020, Jamia Hamdard, New Delhi, India (p. 9).

[3] Bala, R., Gill, B., Smith, D., & Wright, D. (2019). Magic quadrant for cloud infrastructure as a service, worldwide. Gartner Reprint, 1-24.

[4] Malawski, M., Gajek, A., Zima, A., Balis, B., & Figiela, K. (2020). Serverless execution of scientific workflows: Experiments with hyperflow, aws lambda and google cloud functions. Future Generation Computer Systems, 110, 502-514.

[5] Mathur, P. (2024). Cloud computing infrastructure, platforms, and software for scientific research. High Performance Computing in Biomimetics: Modeling, Architecture and Applications, 89-127.

[6] Smith, K. N., & Green, D. K. (2021). Employer internship recruiting on college campuses:'the right pipeline for our funnel'. Journal of Education and Work, 34(4), 572-589.

[7] Bender, D. (2020). Education and Career Skills Acquired during a Design Internship. International Journal of Teaching and Learning in Higher Education, 32(3), 358-366.

[8] Goldstein, H. W., Pulakos, E. D., Passmore, J., & Semedo, C. (2017). The Psychology of Employee Recruitment, Selection and Retention. The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention, 1-11.

[9] Kumar, A. S. (2019). An Exploratory study on use of cloud computing for talent acquisition in health industry. Indian Journal of Public Health Research & Development, 10(8), 221.

[10] Dong, X., & Salwana, E. (2022). The impact of cloud-based human resource and supply chain management systems on the performance of multinational organizations. Kybernetes, 51(6), 2030-2043.

[11] Ahmad, S., Hussain, S., Wasid, M., Onyarin, O. J., Arif, M., & Ahmad, J. (2024). The Future of Recruitment: Using Deep Learning to Build Intelligent Interview Bots. In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE (pp. 1-6).

[12] Pandit, P., Tirlotkar, J., Bait, R., Rao, J., Jha, J., & Deshpande, K. (2024). Comprehensive AI-Driven Web Framework for Automating Recruitment Process of IT Industry. In International Conference on Information and Communication Technology for Intelligent Systems, Singapore: Springer Nature Singapore (pp. 317-326).

[13] Tobroni, M. M., & Hamid, I. R. A. (2023). Secured Recruitment Management System with Data Retention Approach. Applied Information Technology and Computer Science, 4(2), 230-247.

[14] Roy, A., Sural, S., Majumdar, A. K., Vaidya, J., & Atluri, V. (2021). Optimal employee recruitment in organizations under attribute-based access control. ACM Transactions on Management Information Systems (TMIS), 12(1), 1-24.

[15] Jeong, W. Y., & Choi, M. (2019). Design of recruitment management platform using digital certificate on blockchain. Journal of Information Processing Systems, 15(3), 707-716.

[16] Chavan, P. R., Chandurkar, Y., Tidake, A., Lavankar, G., Gaikwad, S., & Chavan, R. (2024). Enhancing recruitment efficiency: An advanced applicant tracking system (ATS). Industrial Management Advances, 2(1), 6373-6373.

[17] Kurek, J., Latkowski, T., Bukowski, M., Świderski, B., Łępicki, M., Baranik, G., ... & Dobrakowski, Ł. (2024). Zero-shot recommendation AI models for efficient job–candidate matching in recruitment process. Applied Sciences, 14(6), 2601.

[18] Alagha, A., Mizouni, R., Singh, S., Otrok, H., & Ouali, A. (2021). SDRS: A stable data-based recruitment system in IoT crowdsensing for localization tasks. Journal of Network and Computer Applications, 177, 102968.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(49s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[19] Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., & Fidler, S. (2015). Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. In Proceedings of the IEEE international conference on computer vision (pp. 19-27).
[20] Chelba, C., Mikolov, T., Schuster, M., Ge, Q., Brants, T., Koehn, P., & Robinson, T. (2013). One billion word benchmark for measuring progress in statistical language modeling. arXiv preprint arXiv:1312.3005, 1-6.
[21] Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). Squad: 100,000+ questions for machine comprehension of text. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (pp. 2383-2392).
[22] Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. R. (2018). Glue: A multi-task benchmark and analysis platform for natural language understanding. In Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP (pp. 353-355).
[23] Williams, A., Nangia, N., & Bowman, S. R. (2017). A broad-coverage challenge corpus for sentence understanding through inference. arXiv preprint arXiv:1704.05426, 1-11.
[24] Chen, Z., Zhang, H., Zhang, X., & Zhao, L. (2018). Quora question pairs.
[25] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 conference on empirical methods in natural language processing (pp. 1631-1642).
[26] Warstadt, A., Singh, A., & Bowman, S. R. (2019). Neural network acceptability judgments. Transactions of the Association for Computational Linguistics, 7, 625-641.
[27] Cer, D., Diab, M., Agirre, E., Lopez-Gazpio, I., & Specia, L. (2017). Semeval-2017 task 1: Semantic textual similarity ultilingual and crosslingual focused evaluation. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), Vancouver, Canada. Association for Computational Linguistics (pp. 1-14).
[28] Dolan, B., & Brockett, C. (2005). Automatically constructing a corpus of sentential paraphrases. In Third international workshop on paraphrasing (IWP2005), 9-16.
[29] Bentivogli, L., Clark, P., Dagan, I., & Giampiccolo, D. (2009). The Fifth PASCAL Recognizing Textual Entailment Challenge. TAC, 7(8), 1.
[30] Levesque, H. J., Davis, E., & Morgenstern, L. (2012). The winograd schema challenge. Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning (pp. 552.561).
[31] Zellers, R., Bisk, Y., Schwartz, R., & Choi, Y. (2018). Swag: A large-scale adversarial dataset for grounded commonsense inference. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), Brussels, Belgium (pp. 93-104).

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.