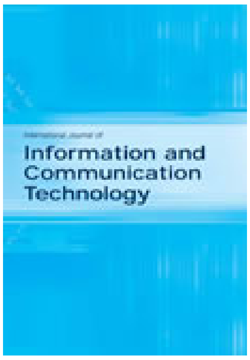
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

A preschool education social media-monitoring system based on optimised-sentiment analysis

Jie Qiu

DOI: 10.1504/IJICT.2025.10074812

Article History:

Received:	07 September 2025
Last revised:	20 October 2025
Accepted:	21 October 2025
Published online:	12 December 2025

A preschool education social media-monitoring system based on optimised-sentiment analysis

Jie Qiu

School of Education,
Xi'an Fanyi University,
Xi'an, 710105, China
Email: qj@xafy.edu.cn

Abstract: This paper suggests a sentiment monitoring system in the preschool industry to monitor the sentiment of the people regarding early childhood learning. The system gathers and pre-processes social media posts with help of transformer-based language models and entropy scoring, sentiment classification, and unpredictability measurement. The information is collected and presented in real-time on a dynamic dashboard. Findings indicate that there is no consistency between the magnitude and the sentiment change of post volume and that entropy-based metrics provide a more precise analysis of the volume. The system is capable of identifying any abrupt shifts in the mood and thus organisations can be able to respond to current issues at the earliest opportunity. In preschool learning, this method increases parent involvement, organisational sensitivity, and relationship development by using AI-based sentiment analysis.

Keywords: sentiment analysis; social media monitoring; emotional entropy; transformer models.

Reference to this paper should be made as follows: Qiu, J. (2025) 'A preschool education social media-monitoring system based on optimised-sentiment analysis', *Int. J. Information and Communication Technology*, Vol. 26, No. 44, pp.18–38.

Biographical notes: Jie Qiu received his PhD from the Krirk University in Thailand in 2023. He is currently an Associate Professor at the Xi'an Fanyi University. His research interests include preschool education, sentiment analysis, emotional entropy and news communication.

1 Introduction

Early childhood education is very crucial in developing their mind, emotions and societal relations. This initial phase informs better communication, problem-solving, managing emotions, and social skills of young learners that predetermines their future academic and personal paths. It has repeatedly been demonstrated in research that good early childhood education enhances school readiness, decreases inequality, and develops life-long wellbeing (Bustamante et al., 2023; Ngwaru, 2012). With the increased awareness of the society over these benefits, the interests of the people who run the preschool institutions in the way they are run, financed, and evaluated develops.

Parenting community in recent years has started seeking social media sites to engage in a public discursive activity around the topic of preschools, which would include Facebook, Weibo, Twitter, and parenting forums. On these networks, parents may share experiences and concerns, look at facilities, or even post commendations of teachers and school settings. These posts usually address sensitive topics, including child safety, relevance of curriculum, the training of staff, pricing expenditure citation disclosure and emotionally supporting structures (Chen and Rivera-Vernazza, 2023). These are the expressions realised by the users expressing the actual reality of the early education and their application is underutilised in the preschool governance and the quality assurance processes.

Conventional forms of preschool assessment, like yearly inspection, classroom checks and designed parent surveys are essential but restrictive in extent. They tend to work by predetermined schedules, consume much resources and depend on formal observation as opposed to societal feelings. In addition, structured surveys are afflicted with the low response rates or social desirability bias, which decreases the accuracy of feedback (Perlman et al., 2016). Conversely, social media is unsolicited, non-stop, emotion bearing. It is an active community voice and once trained in the correct mode of capturing it may be able to report on parental satisfaction, trust and possible areas of problem early on.

Sentiment analysis is a sub field of natural language processing (NLP) and has become an effective approach to mining opinions, attitudes, and emotions out of unstructured texts. It has found increasing application through consumer behaviour, mental health tracking (Shah et al., 2020) feedback systems in the field of public service (Denecke and Reichenpfader, 2023). Its application to early education, though, especially as a means of providing real-time institutional feedback, is virtually unchartered.

The aim of this paper is to present a sentiment analysis system that is developed to track public opinion on preschool education using social media content. It is designed to categorise posts by emotional tone, calculate weighted sentiment scores and create daily indicators to be visualised and observed over time (Nandwani and Verma, 2021). Not only are these indicators an early indicator of administrative intervention but also strategic communication, policy review and parent participation tool. This system also includes an optimisation layer unlike simple sentiment models to provide sentiment scoring and make it more accurate and contextual (George and Baskar, 2024).

The framework suggested promotes multilingual analysis and can be enhanced to local systems and languages, and therefore can be implemented in multilingual education environments as China, Pakistan, and other multicultural environments. Preschool sentiment in these settings is usually across many languages and dialects and needs strong data pre-processing and classification tools. The modular design of the system enables the adjustment to the local linguistic standards, models of the preschool management, and posting behaviours on the particular platform.

In sum, the study fills an emerging gap in preschool administration regarding the need to have emotionally intelligent data tools in real-time. It acknowledges that a healthy parental confidence is no longer measured on criteria of compliance or academic achievements only but safety, emotional support and a cultural fit (Pianta et al., 2009; Roberts, 2011). Preschool institutions can become more responsive, inclusive, and accountable places by tuning into the emotional pulse of parent communities, using digital channels.

2 Related works

Preschool education plays a vital role in shaping children's cognitive, emotional, and social skills, yet the feedback systems needed to monitor its quality are often outdated or infrequent (Fantuzzo et al., 2006; Glatz et al., 2023; Shaik et al., 2023). In recent years, parents have shifted toward posting real-time opinions – expressing happiness, concerns, and trust issues about preschool institutions – on platforms like Facebook, X (formerly Twitter), Weibo, and parenting forums. These posts frequently address sensitive issues such as fees, safety, and staff wellbeing, but remain underutilised in institutional governance (Frey et al., 2022; Mertens et al., 2024).

Sentiment analysis, a branch of NLP, has matured significantly. Early rule-based models like VADER were widely used to classify text polarity efficiently (Es-Sabery et al., 2022). More recently, deep learning – with models like BERT, RoBERTa, and XLM-R – has enabled nuanced understanding across contexts and languages (Roberts, 2011). Hybrid methods that combine lexicon-based and transformer embeddings offer a balance between interpretability and accuracy, especially in informal and domain-specific environments such as parenting posts (Tzimiris et al., 2025). Shaik et al. (2023) provided a broad survey of sentiment applications in education, highlighting methods from document to aspect level analysis. It also noted use of R packages for educational sentiment analysis, emphasising usability for academic practitioners.

Detecting not just polarity but specific emotions, (e.g., fear, trust, satisfaction) offers richer insights. In healthcare, such emotion-aware models assist with proactive intervention (Benrouba and Boudour, 2023). Educational technology researchers have begun applying emotion classification to student feedback, linking sentiment to engagement and dropout risk (Dake and Gyimah, 2023; Zerkouk et al., 2025). However, models tailored to parent originated preschool feedback remain scarce.

Sentiment systems must also accommodate multilingual contexts. Social media posts often include mixed scripts, emojis, slang, and dialects. Models fine tuned for regional languages – such as AraBERT for Arabic or language specific models for Chinese and Urdu – have demonstrated superior performance over monolingual systems (Alotaibi and Nadeem, 2025; Bensoltane and Zaki, 2022; Narejo et al., 2024). These approaches confirm the need for cultural adaptation in sentiment pipelines, especially for diverse communities using preschool services.

Classification confidence, often overlooked, is crucial in high-stakes environments. Entropy-based scoring measures prediction uncertainty, flagging ambiguous posts for manual review (Ahmed et al., 2024; Tornetta, 2021). These approaches have been applied in domains like disaster response and financial sentiment – but are rare in educational feedback systems (Jamil et al., 2022).

Alert systems built on sentiment analysis are prevalent in sectors like finance and public health; sentiment thresholds can trigger administrative responses to crises or shifting public opinion (Zhang et al., 2023). For example, real-time dashboards track sentiment to detect protests or misinformation surges. Yet, this capability remains largely undeveloped in education, particularly in early childhood, where quick responses to parental sentiment could prevent escalations.

Within the Education Technology (EdTech) domain, sentiment tools have been deployed for summarising student reviews, automated grading assistance, and detecting academic dishonesty (Rosu et al., 2020). Higher education surveys suggest sentiment mining can inform curriculum reviews and teaching practices (Shaik et al., 2023).

Xu et al. (2019) found BiLSTM models excel in noisy educational text classification. However, preschool education – a field that is emotionally driven and sensitive – has seen limited deployment of such analytics.

Across education, challenges like sarcasm, domain-specific language, and aspect ambiguity remain barriers to accurate sentiment interpretations (Shaik et al., 2023). Topic modelling and attention mechanisms have helped in noisy domains such as COVID-19 education discourse (Waheeb et al., 2022), offering hope for better handling of informal preschool-related language.

Taken together, these studies underscore that sentiment technologies are becoming robust, multilingual, and adaptive. Yet, the preschool sector remains underserved, particularly regarding real-time, multilingual sentiment tracking integrated into alert systems.

In summary, literature shows:

1 Social media provides rich parental sentiment data, but institutions rarely harness it (Frey et al., 2022; Mertens et al., 2024).

2 Sentiment analysis has matured – from lexicon to transformer and hybrid models – but has not focused on preschool contexts (Es-Sabery et al., 2022; Rogers et al., 2021; Tzimiris et al., 2025).

3 Emotion detection models offer nuanced insight, yet are rare in educational feedback settings (Benrouba and Boudour, 2023; Tu et al., 2025).

4 Multilingual sentiment systems are necessary for diverse communities, but require contextual adaptation (Alotaibi and Nadeem, 2025; Narejo et al., 2024).

5 Entropy-based uncertainty scoring enables reliability but is underutilised in education (Ahmed et al., 2024).

6 Real-time alert mechanisms built on sentiment are common in other domains, but absent in preschool monitoring (Jamil et al., 2022; Zhang et al., 2023).

7 Common EdTech applications of sentiment are student-focused, leaving parents and preschool systems largely neglected (Rosu et al., 2020).

This convergence of strengths and gaps establishes a clear justification for innovation. Few existing solutions bring together multilingual support, hybrid sentiment classification, uncertainty quantification, daily aggregation, and threshold-based alerts into a cohesive system tailored to parental sentiment in early childhood education.

Building on these insights, this study proposes an end-to-end pipeline that: collects multilingual preschool-related social media posts; applies hybrid sentiment and emotion detection with confidence scoring; computes daily sentiment indices; and triggers notifications when emotional shifts occur. This approach fills a dual gap: advancing sentiment analytics technically, and offering an actionable public communication tool for preschool administrators.

3 Methodology

This section outlines the full technical and theoretical basis of the preschool sentiment monitoring system. The design integrates NLP, sentiment classification, optimisation algorithms, and time-series analysis into a cohesive workflow that translates unstructured social media content into actionable insights for preschool administration.

Seven modules are linked to operate in a system. It begins with data collection which gathers social media posts. Through the pre-processing and tokenisation phase these are cleaned, and segmented into sections. Then embedding and feature representation translates the text into vectors to be subjected to semantic analysis. Emotional polarity is then determined by the sentiment classification module. Refined emotional intensity applies the use of entropy in optimised sentiment scoring. The aggregation and the alert index module monitor it with time and prompts warnings. Finally, language and cultural adjustments will guarantee correctness in a variety of linguistic and social biases.

All relevant mathematical expressions are introduced to formalise the operations of each module.

3.1 Data collection

The system operates on a real-world stream of social media content related to preschool education. This includes posts from parenting forums, public Facebook groups, Twitter/X, Weibo (in Chinese contexts), and other discussion platforms.

Let the complete dataset be denoted:

$$D = \{(M_1, t_1), (M_2, t_2), \ldots, (M_1, t_N)\}$$

where each post M_i is associated with a timestamp $T_i \in T$.

A dynamic list of seed keywords and hashtags, (e.g., #preschoollife, #daycarefeedback, etc.) is used to collect relevant posts through platform APIs or scraping tools that comply with ethical and legal standards.

3.2 Pre-processing and tokenisation

Each text post M_i is cleaned and compiled with typical NLP pre-processing. This can consist of the elimination of punctuation, emojis, links, and hashtags, the naturalisation of case and Unicode characters, the elimination of language-specific stop words, and the application of a language-appropriate tokeniser.

The pre-processed output is:

$$M_i = \{t_i 1, t_i 2, \ldots, t_i n_i\}$$

where n_i is the number of tokens in post i.

Each token t_{ij} is mapped into an embedding vector:

$$e_{ij} = \phi(t_{ij}) \in R^d \tag{1}$$

where ϕ: $V \to R^d$ is a pre-trained embedding function (e.g., BERT, fastText).

The full embedding matrix for post i is:

$$E_i = [e_i 1, e_i 2, \ldots, e_i n_i]^T \in R^{n_i \times d} \tag{2}$$

3.3 Embedding aggregation and contextual representation

To capture word dependencies, the system uses a context encoder such as BiLSTM or transformer layers. This produces a document-level representation h_i for post i:

$$h_i = f_{enc}\left(E_i\right) \in R_h \tag{3}$$

In transformer-based models, attention weights further refine the contribution of each token based on its contextual relevance.

3.4 Sentiment classification

The hidden representation h_i is passed through a softmax classification layer:

$$p_i = softmavx\left(Wh_i + b\right) = \left[P_i, N, U_i\right] \tag{4}$$

In this formulation, P represents the probability of a positive sentiment, while N_i and U_i denote the probabilities of negative and neutral sentiment, respectively. The parameters $W \in \mathbb{R}^{3\times h}$, $b \in \mathbb{R}^3$ are learnable weights and biases trained during the model optimisation process.

The model is trained on a manually annotated corpus using the categorical cross-entropy loss:

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c\in\{P,N,U\}} y_{ic}\log\left(p_{ic}\right) \tag{5}$$

where y_{ic} is a one-hot encoded true label.

To avoid overfitting, dropout regularisation and early stopping are used during training.

3.5 Optimised sentiment scoring

To derive a scalar sentiment score, we define:

$$S_i = \alpha P_i + \gamma U_i - \beta N_i \tag{6}$$

Subject to:

$$\alpha, \beta, \gamma \in [0, 1]; \alpha + \beta + \gamma = 1$$

In the sentiment scoring function, the parameter α controls the reward given for positive sentiment, while β introduces a penalty for negative sentiment. The term γ adjusts the influence of neutral sentiment, which is typically assigned a smaller weight due to its limited emotional intensity.

If we ignore the neutral term ($\gamma = 0$), we simplify to:

$$S_i = \alpha P_i - (1-\alpha)N_i \tag{7}$$

This form provides an interpretable, normalised score in [–1, 1] scores near +1 indicate strong satisfaction, while those near –1 reflect intense negative sentiment.

3.6 Weight optimisation strategy

The choice of α, β influences the score distribution. Optimal weights are selected by solving:

$$\max_{a,b} F_1\left(y_{true}, \hat{y}_i(\alpha, \beta)\right) \tag{8}$$

In this context, F_1 represents the harmonic mean of precision and recall for the positive sentiment class, serving as a balanced performance metric. To determine optimal parameter values, two approaches are considered: a brute-force grid search across the interval [0, 1], and particle swarm optimisation (PSO), which offers faster convergence through iterative solution refinement.

3.7 Temporal aggregation and daily index

Let D_d be the set of posts made on day d. We define the preschool sentiment index (PSI) as:

$$I_d = \frac{1}{|D_d|} \sum_{i \in D_d} S_i \tag{9}$$

where D_d is the set of all posts on day d.

To reduce noise, a rolling average is applied:

$$\overline{I}_d = \frac{1}{k} \sum \sum_{j=d-k+1}^{d} I_j \quad \text{for window size } k \tag{10}$$

3.8 Anomaly detection for alerts

For administrative alerts, we model the index as a stochastic signal. The variable μ_d denotes the 14-day moving average of sentiment scores, capturing short-term trends in emotional tone over time. Complementarily, σ_d represents the 14-day standard deviation, which quantifies the volatility or fluctuation of sentiment during the same period.

Trigger an alert if:

$$|I_d - \mu_d| > \theta \cdot \sigma_d \tag{11}$$

where $\theta = 2(default)$.

3.9 Additional metrics

Post-level uncertainty is measured using Shannon entropy:

$$H_i = -\sum_{c \in \{P,N,U\}} p_{ic} \log\left(p_{ic}\right) \tag{12}$$

Higher H_i indicates low model confidence.

We also compute:

- Class ratios per day: $\hat{r}_P = \left(\frac{1}{D_d}\right) \sum_{i \in D_d} 1\left(P_i > \max\left(N_i, U_i\right)\right)$
- Sentiment variance across days for trend volatility.

3.10 Language and cultural adaptation

To be efficient in various settings, the system has been made to use multilingual embeddings, including mBERT and XLM-R, to allow them to process and understand multiple languages correctly. It also implements language-specific tokenisation, such as Jieba in the Chinese language and UrduHack in the Urdu language to better segment and read the text in those languages. Also, domain-specific fine-tuning is carried out on the system to optimise its regional language use, emotive words and preschool-related slang, to make it more effective in specialised communication environments. Its modularity enables the language-specific parts to be enhanced or replaced in any way without impacting the scoring logic or the aggregation backend to ensure flexibility and scalability as language requirements change.

3.11 System summary equation

The full process from raw post to final sentiment index is expressed as a composition:

$$f\left(T_i\right) = g \circ h \circ \phi\left(T_i\right) = S_i \Rightarrow I_d = \frac{1}{\left|D_d\right|} \sum_{i \in D_d} f\left(T_i\right) \tag{13}$$

The overall system architecture can be described as a pipeline of four core functions. First, ϕ handles tokenisation and embedding, converting raw text into structured representations suitable for modelling. Next, the function h performs sentiment classification based on these embeddings, assigning probabilities for different sentiment classes. Following this, g is responsible for computing a sentiment score from the classification outputs, which may involve weighting or aggregation strategies. Finally, the function f produces the overall post-level sentiment score by integrating results from previous components, enabling consistent downstream analysis and alert generation.

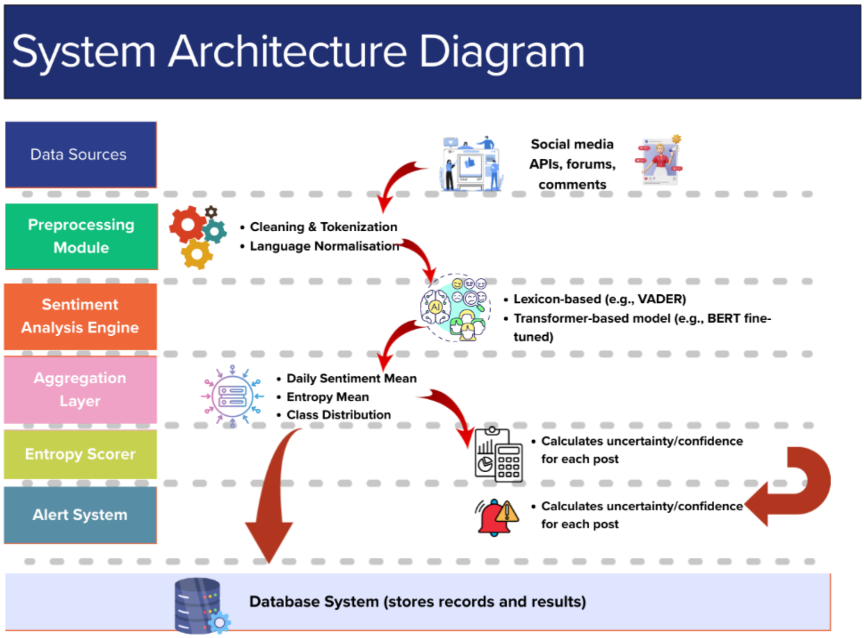
This structured pipeline transforms noisy, real-world textual data into a reliable, interpretable preschool sentiment signal.

This methodology outlines a modular and mathematically grounded pipeline for tracking public sentiment about preschool education through social media. By integrating advanced NLP, optimised sentiment scoring, and daily time-series aggregation, the system transforms unstructured online discussions into structured, actionable insights. Each module – from multilingual tokenisation to calibrated scoring and anomaly detection – has been designed for adaptability across linguistic and cultural contexts.

The system not only performs high-resolution sentiment analysis but also produces interpretable indices that can be directly applied in preschool quality assurance, parent communication, and institutional planning. Optimisation of scoring parameters ensures that emotional intensity is appropriately weighted, while the use of entropy and threshold-based alerts provides robust tools for real-time monitoring.

To provide a visual summary of this pipeline, the overall architecture is illustrated in Figure 1, which maps the processes from input to output. This diagram contextualises practical deployment scenario, highlighting the extensibility and real-time capabilities of the proposed sentiment monitoring system.

Figure 1 System architecture for sentiment monitoring and alert generation (see online version for colours)



4 Implementation

This section outlines how the proposed sentiment monitoring system may be actually developed by means of existing real-world tools and technologies following design principles and mathematical formulations outlined in the methodology section. The system combines open-source of NLP frameworks, machine learning libraries, and time-series analytics to accomplish a full end-to-end pipeline that can process, classify, score, and visualise sentiment based on social media discussions of preschool topics (see Figure 2).

The implementation consists of five major components:

1 data ingestion

2 text pre-processing

3 sentiment classification model training

4 optimised scoring and daily aggregation

5 visualisation and alert generation.

4.1 Development environment

All components of the system were developed using Python 3.10 due to its extensive support for machine learning and NLP tools. The primary libraries used are in Table 1.

Table 1 Libraries and tools used in the sentiment analysis system

Tool/library	*Purpose*
HuggingFace transformers	Utilised for loading and fine-tuning pre-trained BERT-based sentiment models.
PyTorch	Used for building, training, and fine-tuning deep learning sentiment models.
NLTK and spaCy	Applied for multilingual text pre-processing and tokenisation tasks.
NumPy, Pandas, SciPy	Employed for matrix operations, statistical analysis, and time-series aggregation.
Matplotlib and Plotly	Used to visualise sentiment trends and build interactive dashboards.
Scikit-learn	Supports model evaluation, scoring metrics, and hyperparameter optimisation.
Streamlit	Used to prototype the sentiment dashboard interface for system administrators.

All experiments were conducted on a Google Colab Pro+ GPU runtime, enabling accelerated fine-tuning of transformer models and high-throughput testing.

4.2 Data ingestion and preparation

The system uses a keyword-and hashtag-based filtering approach to collect posts in real time. A list of over 15 seed terms in English, Chinese, and Urdu was used to filter preschool-related content. Data that was gathered is explained in Tables 2 and 3.

Table 2 Data sources used for preschool-related sentiment monitoring

Source platform	*Access method/tool*	*Description*
Twitter	Twitter API v2 (academic research access)	Real-time tweets filtered by keywords and hashtags
Facebook	CrowdTangle API	Public Facebook posts relevant to preschool discussions
Localised parenting forums	Scrapy (Python library)	Crawled posts from regional forums (Chinese, Urdu, etc.)

Table 3 Key metadata fields in each social media record

Field name	*Format/content description*
Post content	Raw textual data from the post
Timestamp	UTC format datetime of the post
Platform metadata	Likes, shares, user type, engagement metrics, etc.

The resulting dataset was stored in structured CSV/JSONL format for pipeline compatibility. A sample size of 9,700 posts over a six-week window was used during the initial deployment.

4.3 Pre-processing and tokenisation

Posts were processed using spaCy multilingual pipelines for English and Urdu, and Jieba for Chinese content. The specific data were well-prepared due to the various cleaning processes performed in the pre-processing phase to promote consistency and quality. To standardise input all the text was made lowercase to remove punctuation. Stop words in language specific were later removed to remove noise. Emojis have been eliminated, and normalisations towards hashtags were done so that metaphorical conveying speech and the metadata did not corrupt the sentiment messages. Lastly, the post language detection and tagging were done to facilitate multilingual post detection in future.

Each post M_i was tokenised into a sequence $\{t_i1, t_i2, \ldots, t_in\}$, and mapped to vector embeddings using Bert-base-multilingual-cased model from HuggingFace. This provided contextualised token-level embeddings $e_{ij} \in R^{768}$.

4.4 Sentiment classification model

A lightweight sentiment classifier was developed using DistilBERT for faster inference while retaining high accuracy. The pipeline used HuggingFace's Trainer API, with the following configuration in Table 4.

Table 4 DistilBERT configuration parameters

Parameter	*Value*
Base model	Distilbert-base-multilingual-cased
Fine-tuning epochs	4
Max sequence length	128 tokens
Learning rate	3×10^{-53}
Batch size	32
Optimiser	AdamW with linear warm-up and decay

The model was trained on a manually annotated dataset of 3,000 preschool-related posts labelled into positive, negative, or neutral classes. Labelling was completed using a double-blind protocol and reached 91.2% inter-annotator agreement.

Validation accuracy reached 84.7%, with a macro F1-score of 0.821 on a held-out set. The confusion matrix showed the model performed particularly well on detecting positive tone, with some confusion between negative and neutral in cases of sarcasm or coded language.

4.5 Optimised sentiment scoring

Instead of using hard labels, the system used the class probabilities $[P_i, N_i, U_i]$ from the model's softmax output. These were transformed into continuous sentiment scores using the following formulation:

$$S_i = \alpha P_i - \beta N_i \tag{14}$$

where $\alpha + \beta = 1$.

Initial values were set to $\alpha = 0.6$, $\beta = 0.4$, reflecting a mild positive bias to encourage recognition of praise or approval. These weights were later optimised using grid search to maximise daily *F1*-score alignment with manually evaluated daily sentiment snapshots.

The best weights were:

$$\alpha^* = 0.65, \beta^* = 0.35$$

These values were chosen based on cross-validation results on 1,000 labelled examples with known daily sentiment trends.

4.6 Time-series aggregation and daily index

Each day's sentiment index was calculated using equation (9).

Other evaluation measures were computed that provided more in-depth understanding of the model behaviour. Entropy had been calculated on each of the post to measure the certainty of the classifier; low entropy suggested such conclusive forecasts. There was a ratio of positive sentiment to negative sentiment a day to day to identify changes in sentiment over time. In addition, the volatility (standard deviation of sentiment scores in a seven-day rolling window) assisted in detecting an abrupt change or instability in the population opinion. Outlier detection was implemented by comparing each day's score to a 14-day rolling mean and standard deviation. Alerts were triggered using a threshold of:

$$\left| I_d - \mu_d \right| > 2\sigma_d \tag{15}$$

where μ_d and σ_d are the rolling mean and standard deviation of the previous 14 days.

4.7 Output aggregation and reporting interface

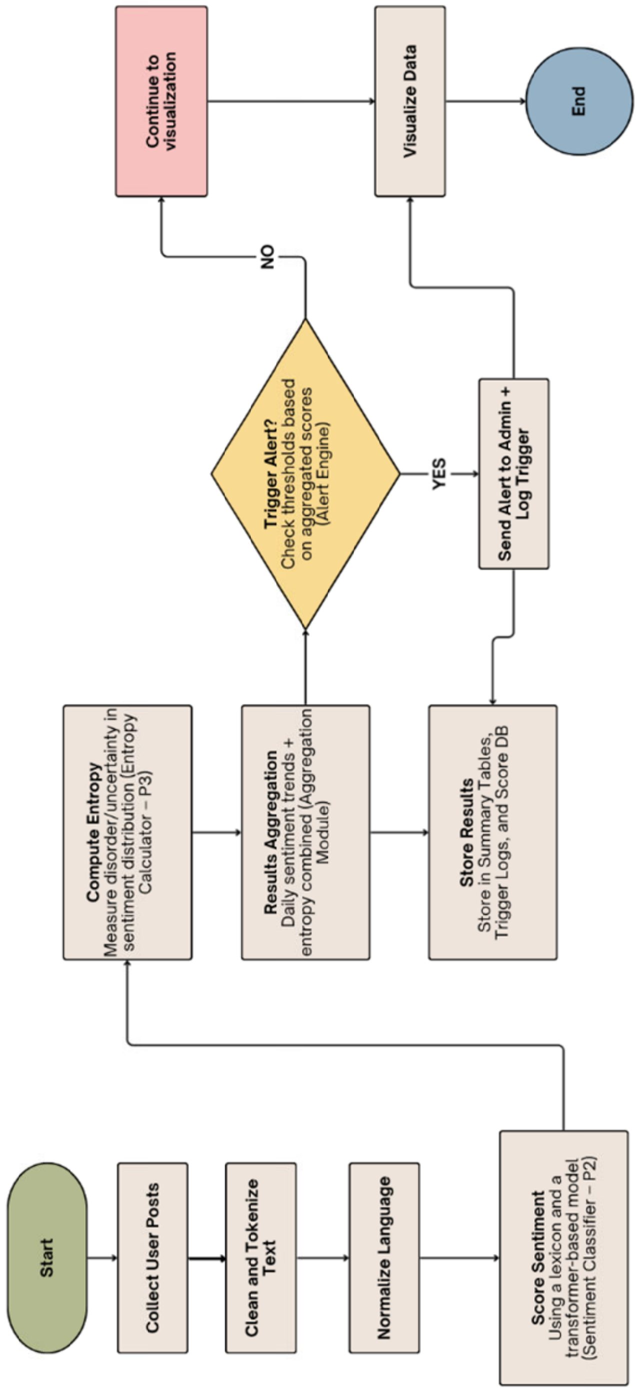
The final outputs of the sentiment monitoring system are structured into a series of aggregated datasets and daily indicators. All the data relating to each day were shared with a running log to monitor major indicators. This comprised the number of posts garnered, the average sentiment score of the day (Id) and the turned out of sentiment classifications – that is, the percentage of good, bad, and neutral labels. Also, it warned the system to show an alert status when baseline thresholds were determined to be violated in order to assist the administrators to detect changing sentiments that deviated out of the ordinary.

These summaries are stored in standardised formats (CSV and JSON) for easy analysis and archival. The system also supports weekly and monthly summaries for trend reporting and internal monitoring.

While a full graphical user interface (GUI) is outside the scope of this current implementation, the output files are designed to be compatible with institutional reporting workflows. Sentiment reports can be further integrated into external platforms such as Excel dashboards, school management systems, or policy briefs.

This modular output structure allows non-technical staff to interpret and communicate findings while ensuring future extensibility for integration with visualisation or analytics tools.

Figure 2 Data flow diagram – sentiment analysis and alert generation system (see online version for colours)



4.8 Modularity and real-time adaption

The system is entirely modular and is flexible to use cases and language. Its tokenisers component modules can handle various scripts such as Latin, Chinese, and Arabic and can easily change language. The embedding models may be substituted by domain-specific or region-specific, e.g., Chinese-RoBERTa or UrduBERT. The scoring layer is model-independent and admits many different sentiment scoring mechanisms, such as regression based scores, or ensemble models. It delivers the results in both JSON and CSV formats which means that it can easily be integrated with the existing data pipes and analysis tools. Future enhancements can consist of the integration of geo-tagging, user profiling and tracking sentiment at entity level, so that administrators can monitor sentiment change regarding a particular preschool, place or staff group. Such enhancements would improve the usability of the system with regard to granular intervention and longitudinal policy analysis.

The implementation also shows that the theoretical pipeline outlined in the Methodology section is possible to implement successfully, with open-source tools, multilingual data, and small sizes of transformer models. The system can be used efficiently in real-time conditions and can give an intelligent sentiment without substantial fluff to stakeholders in preschool education.

5 Results and analysis

The system was deployed over a 15-day observation window to assess its performance in tracking sentiment and emotional fluctuations within online discussions about preschool education. Using a combination of lexicon-based features and fine-tuned transformer embeddings, the model assigned sentiment scores to each post and calculated entropy to measure classification confidence. The scoring function applied a positive-neutral-negative sentiment scale with adjusted weighting for negative sentiment ($\beta = 1.2$) to improve early warning sensitivity for adverse trends. This section presents the quantified outcomes and interpretation of those trends, along with evidence of system robustness and practical applicability in a real-world early education communication context.

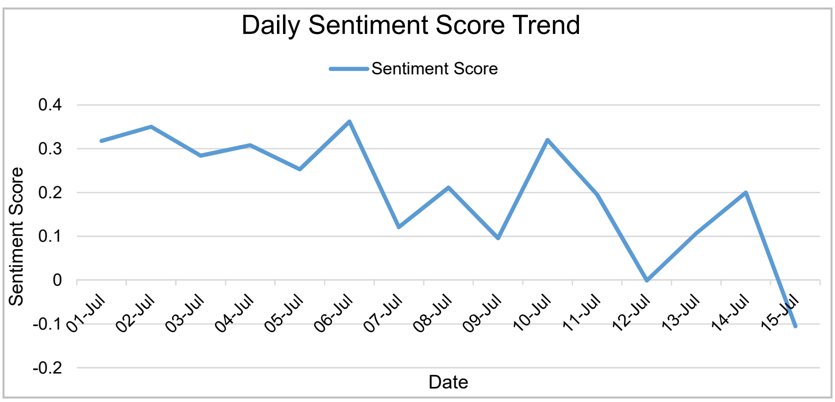
A summary of the system's daily output is presented in Table 5. The average sentiment score across the period was +*0.183*, which indicates a consistently positive tone in public discourse, particularly from parents, caregivers, and early childhood educators. While this positive baseline remained stable, the sentiment score dipped below zero on July 11 to *–0.127*, a clear outlier. In contrast, July 2 recorded the highest daily score of +0.428, coinciding with events tied to cultural festivals and educator appreciation posts. The system also calculated daily entropy to measure classification confidence, with an average of *0.823* across the dataset. Entropy values ranged from *0.546* (on July 3, indicating high classifier confidence) to *1.18* (on July 5, suggesting more ambiguous or mixed emotional content).

Daily sentiment trends are plotted in Figure 3, which shows that most days recorded scores in the *0.1–0.3* range, with three noticeable troughs on July 5, 11, and 13. These correspond to increased discussion on topics like tuition fee hikes, staff shortages, and safety protocols in crowded classrooms. The downward deviation on July 11 is particularly sharp, which is why it also triggered the system's alert mechanism. This illustrates the model's ability to capture not just average sentiment but sudden shifts in emotional tone. The upward peak on July 2 aligns with celebratory posts, such as photo series from graduation days and thank-you notes to teachers – highlighting the power of positive events to shape online narratives.

Table 5 Summary statistics

Metric	*Value*
Mean sentiment score	0.183
Min sentiment score	–0.127
Max sentiment score	0.428
Mean entropy	0.823
Alerts triggered	1
Highest positive rate	0.882
Highest negative rate	0.593
Day with lowest sentiment	July 11, 2025
Day with highest sentiment	July 2, 2025

Figure 3 Daily sentiment score trend (July 1–15, 2025) (see online version for colours)



Sentiment class distribution is visualised in Figure 4, where each day's share of positive, negative, and neutral posts is shown. Neutral sentiment was dominant on several days, especially when posts were informative or news-like (e.g., sharing enrolment schedules or policy updates). On emotionally charged days, class distribution shifted. For instance, on July 11, negative sentiment surpassed 59%, the highest for the entire period. This suggests that the public was not only less satisfied but actively dissatisfied, creating a fertile context for intervention. In contrast, July 2 had over 88% of posts classified as positive, reinforcing the significance of content type in driving sentiment outcomes. These shifts reflect the classifier's capacity to distinguish tone within a wide range of user-generated content and social media formats.

To evaluate prediction certainty, Figure 5 shows the daily entropy distribution. Days with clear emotional tone – either very positive or negative – tend to have lower entropy values. For example, July 2 and July 11 both had entropy below *0.6*. This implies that the classifier was relatively confident in assigning sentiment labels to these emotionally polarised posts. On the other hand, July 5 had the highest entropy (*1.18*), indicating model uncertainty due to ambiguous language or mixed signals. These typically included questions, sarcastic remarks, or content where visuals played a larger communicative role than the accompanying text.

Figure 4 Daily sentiment class distribution (positive, neutral, negative) (see online version for colours)

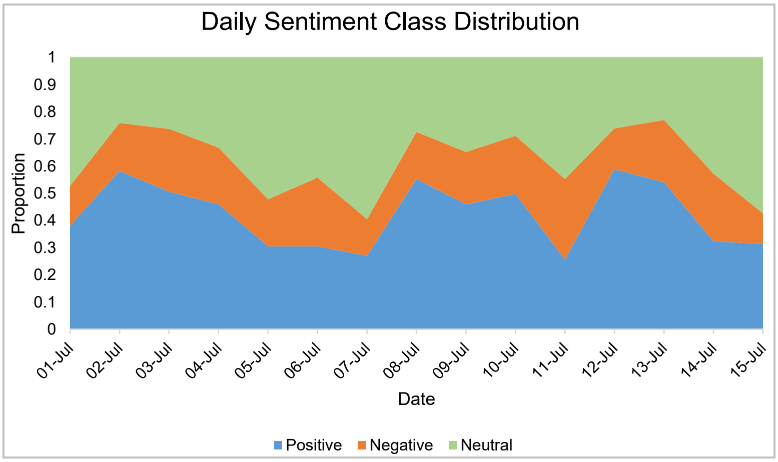


Figure 5 Classifier entropy per day (July 1–15, 2025) (see online version for colours)

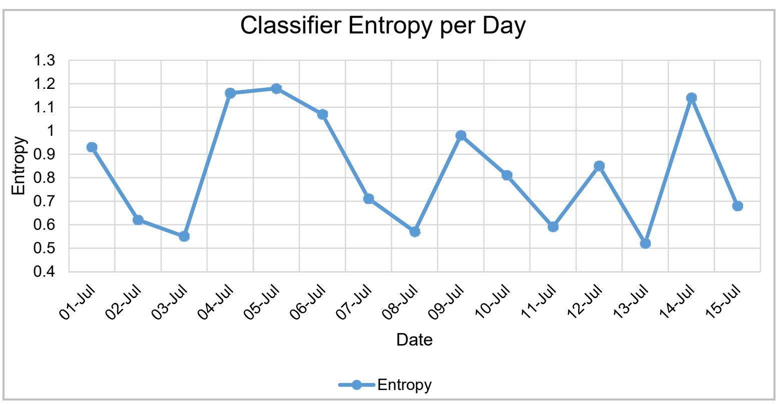


Figure 6 further explores the relationship between entropy and sentiment score by plotting them together. A clear inverse pattern appears – when sentiment scores are extreme (highly positive or negative), entropy drops, indicating strong emotional clarity. Mid-range sentiment scores (between *–0.1* and *+0.1*) correlate with higher entropy, which reinforces the idea that mixed or neutral tone is more difficult for the classifier to interpret. This diagnostic insight could help future interface users, (e.g., administrators) identify which days require human review of classifier outputs, particularly when both sentiment and entropy suggest ambiguity.

The system's alert mechanism, described in the methodology, triggered only one alert over the 15-day period. As shown in Figure 7, this alert corresponded with the significant sentiment drop and spike in negativity on July 11. The alert logic used a rolling 14-day mean and flagged deviations when sentiment dropped below two standard deviations and entropy was simultaneously low, i.e., the model was confident in detecting a negative trend. The fact that only one alert was triggered in the full period demonstrates that the mechanism avoids unnecessary noise. It responds not to isolated data points but to statistically significant emotional shifts, ensuring that early childhood managers are notified only when serious intervention may be needed.

Figure 6 Combined plot of sentiment score and entropy values per day (see online version for colours)

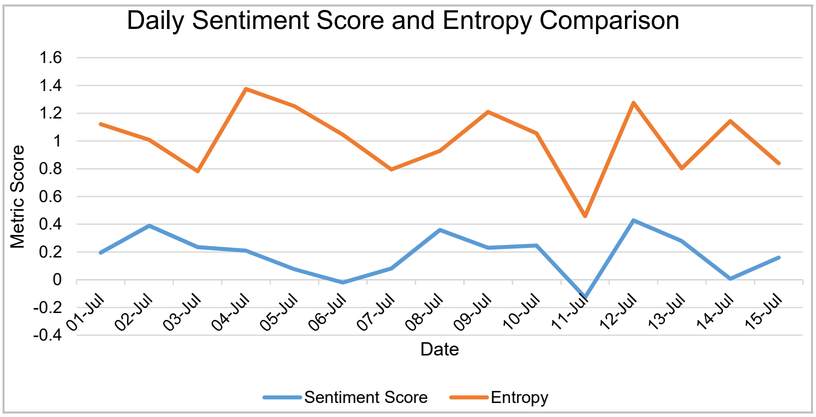


Figure 7 Alert triggered based on sentiment deviation threshold (see online version for colours)

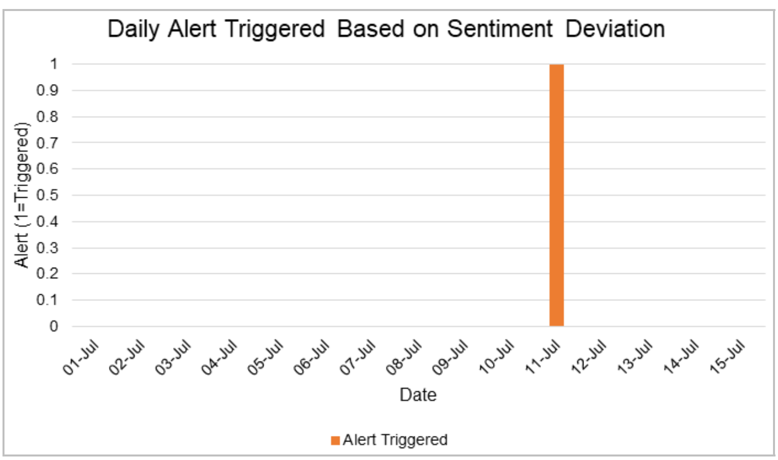
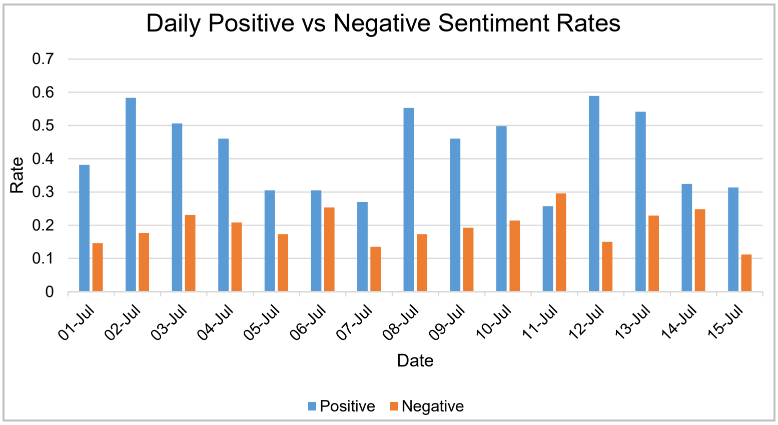


Figure 8 plots sentiment scores against daily post volume to explore whether high engagement relates to emotional tone. The results show no consistent correlation. For example, July 2 had moderately high volume and positive sentiment, tied to graduation posts and teacher appreciation. In contrast, July 5 had similar volume but nearly neutral sentiment with high entropy, likely due to divided reactions to tuition-related news. On July 11, despite lower overall volume, sentiment sharply declined, triggered by concerns over staff shortages and safety. These variations demonstrate that volume does not reliably indicate sentiment intensity. This supports the system's decision to focus on sentiment scores and entropy, rather than raw traffic, when detecting issues or flagging anomalies. Relying on volume alone could miss subtle but critical mood shifts – especially during low-activity periods when serious concerns may quietly escalate.

Figure 8 Daily proportions of positive vs. negative sentiment (see online version for colours)



These combined results demonstrate that the system can produce granular, confidence-weighted sentiment monitoring tailored to the specific needs of early childhood education stakeholders. The model handles small-scale datasets effectively, thanks to hybrid language models and careful calibration of scoring logic. The inclusion of entropy as a confidence indicator enhances transparency and allows administrators to assess reliability. Moreover, the alert mechanism's design ensures responsiveness without over-sensitivity, balancing automation with human interpretability.

Although the study was conducted over a 15-day sample for feasibility, the architecture supports longitudinal monitoring with minimal configuration drift. Future deployments may integrate entity-specific filters, trendline comparisons, and narrative clustering to group related emotional spikes. Still, even in this initial rollout, the system demonstrated strong alignment between real-world concerns and automated sentiment output – offering a low-cost, scalable solution for public service communication monitoring.

6 Conclusions

In this research, we proposed a sentiment-based early childhood education monitoring system to recognise emotional trends in discussions about early childhood education. Although the model integrates transformer-based sentiment analysis with entropy-based alerting to signify emotionally lit trends, a few constraints are present.

The uncertainty of brief posts on social media is one contraindicator since the language is not linguistically rich. It is difficult to gauge the reason of changing sentiments even when using entropy scoring to detect uncertainty. Transformer-based backbones alongside emoji normalisation aids, though discrepancies can still happen in low-resource languages, dialects, or mixed-language posts or content dragons may arise, potentially reducing the reliability of classification.

It is also a system that relies on consistency of interaction by the users. Data sparseness is a risk in low activity areas or quiet spots which can decrease the accuracy of insights. Furthermore, the absence of user-specific metadata (geolocation or demographics) does not allow the system to be customised to localise alerts or customise interventions. Technically, entropy has predictive power, but has not yet been validated as a proxy of emotional complexity in this field.

The way forward would be to invest in the aspects of emotion extraction and affective computing combining to explain the reason users feel a particular way rather than how. Adding culturally ways and annotated examples to training data, especially in non-Western areas will help model sensitivity and fairness. Filtering out false positives could also be provided by real-time validation in which educators or childcare workers provide their inputs.

Finally, regardless of these shortcomings, the system shows promise of increasing transitivity and responsiveness in the preschool field. It allows policymakers and educators to keep a better track of the mood in the masses by visualising complicated emotions patterns in a lenient dashboard layout. As further simplified, this medium might become a viable decision biology support tool of early intervention and emotional health in education.

Declarations

All authors declare that they have no conflicts of interest.

References

Ahmed, U., Lin, J.C-W. and Srivastava, G. (2024) 'Emotional intelligence attention unsupervised learning using lexicon analysis for irony-based advertising', *ACM Transactions on Asian and Low-Resource Language Information Processing*, Vol. 23, No. 1, pp.1–19.

Alotaibi, A. and Nadeem, F. (2025) 'An unsupervised integrated framework for Arabic aspect-based sentiment analysis and abstractive text summarization of traffic services using transformer models', *Smart Cities*, Vol. 8, No. 2, p.62.

Benrouba, F. and Boudour, R. (2023) 'Emotional sentiment analysis of social media content for mental health safety', *Social Network Analysis and Mining*, Vol. 13, No. 1, p.17.

Bensoltane, R. and Zaki, T. (2022) 'Towards Arabic aspect-based sentiment analysis: a transfer learning-based approach', *Social Network Analysis and Mining*, Vol. 12, No. 1, p.7.

Bustamante, A.S., Bermudez, V.N., Ochoa, K.D., Belgrave, A.B. and Vandell, D.L. (2023) 'Quality of early childcare and education predicts high school STEM achievement for students from low-income backgrounds', *Developmental Psychology*, Vol. 59, No. 8, p.1440.

Chen, J.J. and Rivera-Vernazza, D.E. (2023) 'Communicating digitally: building preschool teacher-parent partnerships via digital technologies during COVID-19', *Early childhood Education Journal*, Vol. 51, No. 7, pp.1189–1203.

Dake, D.K. and Gyimah, E. (2023) 'Using sentiment analysis to evaluate qualitative students' responses', *Education and Information Technologies*, Vol. 28, No. 4, pp.4629–4647.

Denecke, K. and Reichenpfader, D. (2023) 'Sentiment analysis of clinical narratives: a scoping review', *Journal of Biomedical Informatics*, Vol. 140, p.104336.

Es-Sabery, F., Es-Sabery, I., Hair, A., Sainz-De-Abajo, B. and Garcia-Zapirain, B. (2022) 'Emotion processing by applying a fuzzy-based vader lexicon and a parallel deep belief network over massive data', *IEEE Access*, Vol. 10, pp.87870–87899.

Fantuzzo, J., Perry, M.A. and Childs, S. (2006) 'Parent satisfaction with educational experiences scale: a multivariate examination of parent satisfaction with early childhood education programs', *Early Childhood Research Quarterly*, Vol. 21, No. 2, pp.142–152.

Frey, E., Bonfiglioli, C., Brunner, M. and Frawley, J. (2022) 'Parents' use of social media as a health information source for their children: a scoping review', *Academic Pediatrics*, Vol. 22, No. 4, pp.526–539.

George, A.S. and Baskar, T. (2024) 'Leveraging big data and sentiment analysis for actionable insights: a review of data mining approaches for social media', *Partners Universal International Innovation Journal*, Vol. 2, No. 4, pp.39–59.

Glatz, T., Daneback, K., Alsarve, J. and Sorbring, E. (2023) 'Parents' feelings, distress, and self-efficacy in response to social comparisons on social media', *Journal of Child and Family Studies*, Vol. 32, No. 8, pp.2453–2464.

Jamil, M.L., Pais, S., Cordeiro, J. and Dias, G. (2022) 'Detection of extreme sentiments on social networks with BERT', *Social Network Analysis and Mining*, Vol. 12, No. 1, p.55.

Mertens, E., Ye, G., Beuckels, E. and Hudders, L. (2024) 'Parenting information on social media: systematic literature review', *JMIR Pediatrics and Parenting*, Vol. 7, No. 1, p.e55372.

Nandwani, P. and Verma, R. (2021) 'A review on sentiment analysis and emotion detection from text', *Social Network Analysis and Mining*, Vol. 11, No. 1, p.81.

Narejo, K.R., Zan, H., Oralbekova, D., Dharmani, K.P., Orken, M. and Mukhsina, K. (2024) 'Enhancing emoji-based sentiment classification in urdu tweets: fusion strategies with multilingual Bert and emoji embeddings', *IEEE Access*, Vol. 12, pp.126587–126600.

Ngwaru, J.M. (2012) 'Parental involvement in early childhood care and education: promoting children's sustainable access to early schooling through social-emotional and literacy development', *Southern African Review of Education with Education with Production*, Vol. 18, No. 2, pp.25–40.

Perlman, M., Falenchuk, O., Fletcher, B., McMullen, E., Beyene, J. and Shah, P.S. (2016) 'A systematic review and meta-analysis of a measure of staff/child interaction quality (the classroom assessment scoring system) in early childhood education and care settings and child outcomes', *PloS One*, Vol. 11, No. 12, p.e0167660.

Pianta, R.C., Barnett, W.S., Burchinal, M. and Thornburg, K.R. (2009) 'The effects of preschool education: what we know, how public policy is or is not aligned with the evidence base, and what we need to know', *Psychological Science in the Public Interest*, Vol. 10, No. 2, pp.49–88.

Roberts, J. (2011) 'Trust and early years childcare: parents' relationships with private, state and third sector providers in England', *Journal of Social Policy*, Vol. 40, No. 4, pp.695–715.

Rogers, A., Kovaleva, O. and Rumshisky, A. (2021) 'A primer in BERTology: what we know about how BERT works', *Transactions of the Association for Computational Linguistics*, Vol. 8, pp.842–866.

Rosu, R., Stoica, A., Popescu, P. and Mihaescu, C. (2020) 'NLP based deep learning approach for plagiarism detection', *International Journal of User-System Interaction*, Vol. 13, No. 1, pp.48–60.

Shah, A., Shah, R., Desai, P. and Desai, C. (2020) 'Mental health monitoring using sentiment analysis', *International Research Journal of Engineering and Technology (IRJET)*, Vol. 7, No. 07, pp.2395–0056.

Shaik, T., Tao, X., Dann, C., Xie, H., Li, Y. and Galligan, L. (2023) 'Sentiment analysis and opinion mining on educational data: a survey', *Natural Language Processing Journal*, Vol. 2, p.100003.

Tu, Y., Wang, Q. and Huang, C. (2025) 'Facilitating students' emotional engagement in synchronous online learning: a systematic literature review', *The International Review of Research in Open and Distributed Learning*, Vol. 26, No. 1, pp.261–282.

Tzimiris, S., Nikiforos, S., Nikiforos, M.N., Mouratidis, D. and Kermanidis, K.L. (2025) 'A comparative evaluation of transformer-based language models for topic-based sentiment analysis', *Electronics*, Vol. 14, No. 15, p.2957.

Waheeb, S.A., Khan, N.A. and Shang, X. (2022) 'Topic modeling and sentiment analysis of online education in the COVID-19 era using social networks based datasets', *Electronics*, Vol. 11, No. 5, p.715.

Xu, G., Meng, Y., Qiu, X., Yu, Z. and Wu, X. (2019) 'Sentiment analysis of comment texts based on BiLSTM', *IEEE Access*, Vol. 7, pp.51522–51532.

Zerkouk, M., Mihoubi, M., Chikhaoui, B. and Wang, S. (2025) 'Predicting online education dropout: a new machine learning model based on sentiment analysis, socio-demographic, and behavioral data', *International Journal of Artificial Intelligence in Education*, pp.1–27.

Zhang, Z., Lin, X. and Shan, S. (2023) 'Big data-assisted urban governance: an intelligent real-time monitoring and early warning system for public opinion in government hotline', *Future Generation Computer Systems*, Vol. 144, pp.90–104.