

IJASC 23-9-7

## **Design and Implementation of an LLM system to Improve Response Time for SMEs Technology Credit Evaluation**

Sungwook Yoon

*Division of SW Human Resources Development, Andong University, Instructor, Korea  
uvgotmail@nate.com*

### **Abstract**

*This study focuses on the design of a GPT-based system for relatively rapid technology credit assessment of SMEs. This system addresses the limitations of traditional time-consuming evaluation methods and proposes a GPT-based model to comprehensively evaluate the technological capabilities of SMEs. This model fine-tunes the GPT model to perform fast technical credit assessment on SME-specific text data. Also, It presents a system that automates technical credit evaluation of SMEs using GPT and LLM-based chatbot technology. This system relatively shortens the time required for technology credit evaluation of small and medium-sized enterprises compared to existing methods. This model quickly assesses the reliability of the technology in terms of usability of the base model.*

**Keywords:** *GPT-based, Small and Medium Enterprises (SMEs), Technical Credit Evaluation, LLM-based Chatbot System, Improve Response Time*

### **1. Introduction**

Small and medium-sized enterprises (SMEs) are vital for economic growth and industrial diversity, but they often face technological limitations due to resource constraints. The need for accurate & rapidly technological credit assessment has grown as traditional methods, which focus on financial status or management, don't fully capture SMEs' actual technological capabilities[1]. To this end, there is increasing interest in an immediate technology evaluation system based on GPT (Generative Pre-trained Transformer) that quickly and relatively accurately evaluates the technology level of small and medium-sized enterprises. GPT is a type of artificial intelligence model developed by OpenAI. GPT models are designed to understand and generate human-like text based on the input they receive. These models have been trained on vast amounts of text data and can generate coherent and contextually relevant responses[2].

GPT models, such as GPT-3, have been widely used in various applications, including natural language

---

Manuscript Received: June. 26, 2023 / Revised: July. 4, 2023 / Accepted: July. 7, 2023

Corresponding Author: uvgotmail@nate.com

Tel:\*\*\* - \*\*\*\* - \*\*\*\*

Division of SW Human Resources Development, Andong University, Professor

understanding, language translation, content generation, and even assisting with tasks like coding or answering questions. They employ deep learning techniques, particularly transformer architectures, which enable them to capture complex patterns and dependencies in the data they are trained on.

In this study, I intend to design a GPT-based fine-tuning system so that the technology credit evaluation of SMEs can be evaluated relatively quickly. This aims to solve the existing evaluation limitations of the existing time-constrained evaluation and propose a GPT-based model that comprehensively evaluates the technological capabilities of SMEs. This system creates a professional credit rating model that takes into account factors specialized in small and medium-sized enterprises. By fine-tuning the GPT model, customized SME text data is learned to perform relatively fast technology credit evaluation.

This study focuses on designing a GPT-based technology credit rating system for SMEs. This starts with understanding the GPT model and exploring its fine-tuning for technology evaluation for small business owners. In this study, a GPT-based credit rating model is designed and implemented to evaluate its performance and compare it with existing methods to demonstrate time efficiency.

## **2. Related Studies**

Credit evaluation is a critical process for assessing the credit risk of corporations or individuals. In the context of SMEs, such assessments play an essential role in investment and fundraising. Credit evaluation methodologies can be broadly categorized into two main types. Traditional Credit Evaluation Models includes multivariate statistical analysis and credit scoring card systems[3]. While they offer validated methodologies, they suffer from the drawback of struggling to capture diverse patterns and complex structures. Multivariate statistical analysis simultaneously analyzes multiple variables to derive credit scores, and credit scoring card systems evaluate by assigning weights to specific variables[4].

Machine Learning-based Credit Evaluation, This methodology encompasses techniques such as Support Vector Machine (SVM) and Random Forest. SVM is capable of capturing nonlinear patterns, and Random Forest is advantageous for complex pattern analysis and overcoming overfitting issues. While machine learning methodologies are favorable for intricate pattern recognition and accuracy enhancement, they might present challenges in interpretability and complexity of the model structure[5]. The comparison and analysis of these credit evaluation methodologies play a vital role in constructing an optimal assessment system aligned with the characteristics and demands of SMEs. This area is considered to need continuous research and development[6].

Existing credit evaluation approaches for SMEs face limitations that hinder accuracy and efficiency[7]. Some models struggle with quantification, leading to subjective judgment and difficulties in handling variables that are hard to quantify. Data sensitivity can also be problematic, with models overly responsive to specific data, reducing generalization. These challenges underscore the ongoing need for research to address these issues.

Credit evaluation is crucial for SME financial decisions, but current methodologies exhibit shortcomings. Traditional models might oversimplify by relying on formalized variables, failing to capture market dynamics or unique contexts[8]. Machine learning methods, while adept at complex patterns, lack interpretability and demand substantial, high-quality data. Efforts to address these problems involve integrating traditional and machine learning models and introducing new, non-formal variables like market trends. Ongoing research and improvement in this complex field are anticipated to bolster SME growth and financial market stability[9].

### 3. Proposing a GPT-Based SME Technology Credit Evaluation System

The Large Language Model (LLM), such as GPT-3, has demonstrated strong performance in various natural language processing tasks, including language understanding, generation, translation, and more. Its ability to process and generate coherent text makes it a valuable tool for applications such as chatbots, content generation, and even assisting with specific tasks.

The GPT model, renowned for strong performance in natural language tasks, is harnessed for SME technology credit evaluation. Built on the Transformer architecture, GPT learns from vast text data, adaptable to diverse NLP tasks, including assessing SMEs' tech abilities and creditworthiness.

Requirements analysis is pivotal, establishing system functions, standards, and limitations. Functional needs span credit assessment automation, user interaction, data management, security, compliance, and customized reporting. Performance expectations cover response time, accuracy, scalability, reliability, and user-friendliness. Constraints consider budget, technology, legalities, and regulations[10]. This analysis aligns project teams, stakeholders, and users, guiding design and implementation for an efficient, accurate SME credit evaluation system.

The architecture of the GPT LLM chatbot system for SME technology credit evaluation can be broadly divided into three parts: the User Interface (UI), chatbot engine, and data & analysis layer. The UI, where direct interaction with users occurs, takes the form of a web or mobile application. The chatbot engine, integrating GPT and LLM, is the core segment responding to user queries and analyzing the necessary information to conduct credit evaluation. The data and analysis layer store and process information on SMEs[11]. Specifically, the UI layer can be subdivided into frontend and backend, the chatbot engine layer into NLP processor, dialogue manager, and credit evaluation module, and the data & analysis layer into database and analytical tools[12].

The system can be broadly categorized into three principal components: the UI, chatbot engine, and data & analysis layer. The UI facilitates direct interaction with users and takes the form of a web or mobile application, serving as the gateway for user communication. The chatbot engine is the core segment that integrates GPT and LLM to respond to user queries, analyze the necessary information, and perform credit evaluations. Its functions are central to the system's ability to interpret and process user requests efficiently. Situated within the data & analysis layer are the database and analytical tools that store and process information concerning SMEs.

This layer plays a vital role in the management and utilization of data, forming the backbone of the credit evaluation process. Together, these components create a comprehensive architecture that defines the system's functionality and flow, ensuring that it operates cohesively to fulfill its intended role in SME technology credit evaluation.

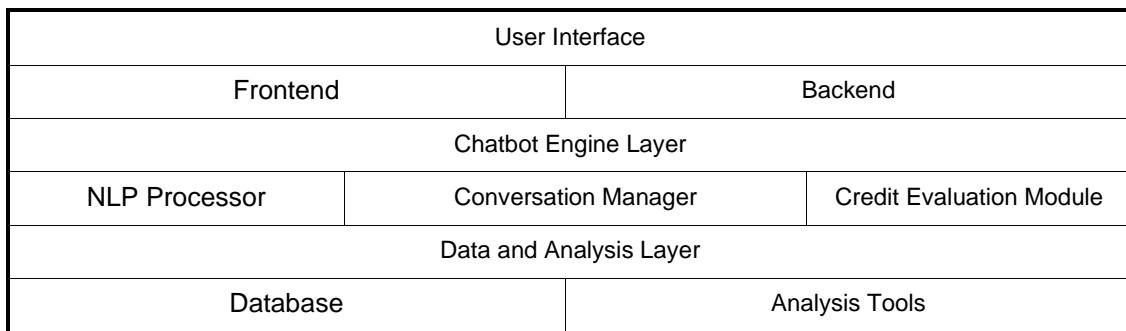
The detailed structure can further be described in terms of specific components and functionalities, such as the following(Table1):

**Table 1. Specific components and functionalities**

Layer	Component	Description
UI Layer	Frontend	Accepts user input and provides output

	Backend	Transfers and processes user requests to the chatbot engine
Chatbot Engine Layer	NLP Processor	Natural Language Processing and understanding
	Dialogue Manager	Controls the flow of conversation
	Credit Evaluation Module	Executes credit evaluation algorithms
Data & Analysis Layer	Database	Stores and manages enterprise data
	Analysis Tools	Supports data analysis and evaluation

The diagram that visually represents the structure of the system clearly illustrates the individual components and their interactions with one another. It can be described as follows(Figure. 1):



**Figure 1. System architecture**

The selection of primary algorithms and technologies employed in the GPT LLM chatbot system for SME technology credit evaluation is articulated in this section. In this system, the GPT and LLM are utilized as key components, with a detailed explanation of the rationale for their selection and methods of application. GPT is chosen for its superior performance in natural language processing, allowing for precise responses to user inquiries. As a pre-trained model, it aids in conserving development time and enhancing generalization performance. Through user interaction handling and customization, it reflects specific domain knowledge in credit evaluation.

The GPT model can be adjusted for specific tasks by fine-tuning the pre-trained model. To fine-tune the GPT model for the technology credit evaluation of SMEs, it must be trained using a relevant dataset. During the fine-tuning process, the input and output of the training data are defined, and hyperparameters are set to optimize the model. Through this, the GPT model can be specialized to suit the technology credit evaluation of SMEs.

The GPT-based SME technology credit evaluation model is designed to utilize the GPT model in performing technology capability assessments of SMEs. This model integrates various aspects of SMEs, considering technology capability, management ability, financial condition, etc., into the ganzfeld technology credit evaluation model. By utilizing the results of the GPT model's fine-tuning, the model takes the SME's text data as input to perform technology credit evaluations, aiming to assess the technical trustworthiness of the enterprise by considering essential evaluation factors. Through this, a system is designed that can provide accurate and reliable technology credit evaluations for SMEs.

## 4. Implementation Results and Analysis

To refine the GPT model for SME technology credit evaluation, In this study compiled a comprehensive dataset aligned with the TCB (Technology Credit Evaluation) criteria. This dataset encompasses an array of attributes crucial for assessing SMEs' technological prowess, financial stability, managerial competence, and market competitiveness. It comprises a diverse collection of samples spanning various industries and sectors, encompassing SMEs engaged in technology-centric domains such as software development, manufacturing, biotechnology, and renewable energy solutions. This diversity has been strategically incorporated to ensure that the trained model can effectively evaluate technological credit across a spectrum of domains.

Each sample within the dataset encompass financial metrics such as revenue, profit margins, assets, and liabilities; technological innovation indicators like research and development investments, patents either filed or owned by the company; market competitiveness markers, including market share analysis, customer satisfaction ratings, and brand awareness metrics; HR and organizational structure data such as employee distribution by department or position, along with assessments of organizational culture; external environmental factors such as industry trend analysis and compliance with legal regulations.

Datasets are meticulously curated, balancing representations and maintaining rigorous data quality standards across different types of small and medium-sized enterprises. A dataset of 50 samples was used to ensure a large amount of training data for effectively fine-tuning the GPT model. Using these diverse and comprehensive datasets in the fine-tuning process, the goal of this study was to improve the model's ability to accurately assess the technology creditworthiness of SMEs to meet TCB criteria. Implementation was carried out to evaluate the performance of the proposed GPT-based SME technology credit rating system. The implementation design took into account the following factors:

- Selection and collection of the SME dataset to be used in the implementation
- Formation of the implementation groups to compare existing technology credit evaluation methods
- Selection of appropriate evaluation metrics for testing
- Setting of implementation conditions and hyperparameters

The implementation of the SME technology credit evaluation GPT LLM chatbot system involves complex natural language processing and data analysis tasks. Consequently, the selection of an effective development environment and tools is essential. The following section describes the main languages, frameworks, libraries, etc., used in the implementation. Below is an example of a Python script responsible for an essential part of the GPT LLM chatbot system for SME technology credit evaluation. The Pseudocode code utilizing the GPT model to respond to user queries is also provided below(Figure 2).

```
IMPORT GPT2LMHeadModel and GPT2Tokenizer from transformers

INITIALIZE model by LOADING GPT2LMHeadModel from "gpt2" pretrained model
INITIALIZE tokenizer by LOADING GPT2Tokenizer from "gpt2" pretrained model

DEFINE FUNCTION generate_response with parameter prompt:
    TOKENIZE prompt using tokenizer and CONVERT to tensor, STORE in inputs
    GENERATE response from model using inputs, SET max_length to 100 and num_return_sequences to 1, STORE in
outputs
    DECODE outputs to get the response text, SKIP special tokens, STORE in response
    RETURN response

PROMPT user to ask a question, STORE in user_prompt
```

```
CALL generate_response with user_prompt, STORE result in response
PRINT response
```

### Figure 2. Pseudocode code utilizing the GPT model to respond to user queries

The project's framework and libraries comprise several key components. For backend development, the Python web framework Django is utilized, responsible for data processing and API implementation(Figure 3). The construction of the user interface employs the JavaScript library React. A significant library included is Hugging Face's Transformers, which is used for training and inference with the GPT model, particularly for fine-tuning the GPT-2 model to specific domains such as SME technology credit evaluation. Figure 2 illustrates the process of fine-tuning the GPT-2 model using SME credit evaluation data.

```
from transformers import GPT2LMHeadModel, GPT2Tokenizer,
TextDataset, DataCollatorForLanguageModeling
from transformers import Trainer, TrainingArguments

# Load the model and tokenizer
model = GPT2LMHeadModel.from_pretrained("gpt2")
tokenizer = GPT2Tokenizer.from_pretrained("gpt2")

# Prepare training dataset
def load_dataset(file_path):
dataset = TextDataset(
tokenizer=tokenizer,
file_path=file_path,
block_size=128
)
return dataset

train_dataset = load_dataset('train_data.txt')

# Set up data collator for preprocessing
data_collator = DataCollatorForLanguageModeling(
tokenizer=tokenizer, mlm=False
)

# Configure training settings
training_args = TrainingArguments(
output_dir="./output",
overwrite_output_dir=True,
num_train_epochs=3,
per_device_train_batch_size=32,
save_steps=10_000,
save_total_limit=2,
)

# Start training process
trainer = Trainer(
model=model,
args=training_args,
data_collator=data_collator,
train_dataset=train_dataset
)

trainer.train()
```

### Figure 3. Data processing code

By utilizing the capabilities of the Transformers library, the training dataset is prepared, training settings are defined, and the model is trained. Fine-tuning is the process of further training an already learned model to suit a specific domain or task. It plays a vital role in enhancing the quality of chatbot responses related to SME credit evaluation.

The implementation results were evaluated through quantitative metrics and statistical analysis. The proposed GPT-based SME technical credit rating system was compared and analyzed against existing credit rating methods. Various evaluation metrics were employed to assess the accuracy, predictability, and reliability of the SME's technical credit rating. Additionally, the characteristics of the SME dataset were taken into account for a specific analysis of the implementation results.

## 5. Evaluation and Verification

System performance evaluation is crucial, determining functionality, goal achievement, and enhancement insights. In the SME technical credit assessment GPT LLM chatbot system, various evaluation methods apply. The evaluation phase prioritizes comprehensive assessment, measuring overall performance via accuracy, scalability, efficiency, and user satisfaction. Identifying system strengths and weaknesses, fostering continuous improvement, is central. This evaluation builds a system with practical problem-solving. Applying tech advancements to SME credit evaluation boosts efficiency and accuracy. Domain training, system integration, security, and compliance ensure success, aiding SMEs with time-saving, resource conservation, and informed credit decisions(Table 2).

The application of the GPT LLM chatbot system for the technical credit evaluation of actual SMEs is illustrated in a hypothetical case study involving a generic company. This typical company is involved in the development of new energy-efficient solutions for smart devices, and their existing manual-based credit evaluation process was both time-consuming and inefficient. The introduction of an automated credit evaluation system aimed to reduce evaluation time and enhance accuracy.

**Table 2. Evaluation metrics and processes for the SMEs technical credit rating system**

Evaluation Category	Metrics	Process
Accuracy Evaluation	RMSE, MAE, and other numerical analysis metrics	Cross-validation, splitting the test set, etc.
Natural Language Response Evaluation	Scoring appropriateness, naturalness, grammatical accuracy, etc.	Presenting predefined questions and system responses to evaluators and requesting evaluation
Scalability & Efficiency Evaluation	Response time, throughput, resource utilization, etc.	Using load testing tools to create virtual users, monitoring performance
User Satisfaction Evaluation	User satisfaction scores collected through surveys and interviews	Providing surveys to system users or conducting interviews

During the implementation process, various data were collected, including the company's financial data, technological innovation metrics, and market competitiveness. Moreover, specific data required for the credit evaluation by the GPT LLM chatbot system were gathered within this generic company, encapsulating the complete process necessary for the new evaluation method(Table 3).

In the study, the evaluation criteria for SMEs are defined in the form of a dictionary, and each category and its sub-items are presented. In an actual evaluation system, these criteria can be used as a basis for data collection, analysis, and assessment(Figure 4).

**Table 3. Key data Categories for SME technical credit evaluation system**

Category	Sub-Items	Description
Financial Data	Revenue, Profit, Assets & Liabilities	Annual/Quarterly revenue, net profit, liquid/fixed assets, long-term/short-term liabilities, etc.

Technological Innovation Indicators	Research & Development Investment, Patents, Product Development	Research investment amount, number of owned patents, product line updates, etc.
Market Competitiveness Indicators	Market Share, Customer Satisfaction, Brand Awareness	Major market share, customer survey results, brand awareness analysis, etc.
HR & Organizational Structure	Number of Employees, Personnel Composition, Organizational Culture	Department/Position-wise employee count, technical/sales staff ratio, organizational culture assessment, etc.
External Environment & Regulations	Industry Trends, Legal & Regulatory Compliance	Analysis of major trends within the industry, compliance with relevant laws and regulations, etc.

Initialization:

Define the evaluation criteria for SMEs as a dictionary

Begin:

Create SME evaluation criteria dictionary

Under "Financial Data", include "Revenue", "Profit", "Assets", "Liabilities"

Under "Technological Innovation Indicators", include "Research & Development

Investment", "Patents", "Product Development"

Under "Market Competitiveness Indicators", include "Market Share", "Customer

Satisfaction", "Brand Awareness"

Under "HR & Organizational Structure", include "Number of Employees", "Personnel

Composition", "Organizational Culture"

Under "External Environment & Regulations", include "Industry Trends", "Legal &

Regulatory Compliance"

End

#### Figure 4. Evaluation criteria

The data collected for credit evaluation at a generic company spans diverse categories including financial data, technology innovation metrics, market competitiveness metrics, HR and organizational structure data, and external environment and regulatory compliance data. These data points collectively enrich SME credit evaluations, enabling a comprehensive assessment of a company's potential from various angles. By integrating these datasets, the business system was streamlined with chatbot automation, allowing for efficient credit evaluation requests and responses. The system's performance and feasibility were tested and verified within real-world business settings, demonstrating its practical applicability (Table 4).

**Table 4. Definition and Functionality of Database in the Fine-Tuning Process**

Component	Definition	Function
Database Schema	Defines database structure, tables, fields, relationships	Supports organized storage and retrieval of fine-tuning data
Data Sources	List and structure of used data sources	Enables diverse data ingestion for model fine-tuning
Data	Techniques to clean, transform,	Ensures data quality and consistency



Preprocessing	normalize data	
Fine-Tuning Algorithm	Specific method to adapt the model	Affects model's adaptation to tasks
Evaluation Metrics	Metrics to assess fine-tuned model performance	Quantifies success in fine-tuning
Access Controls	Security protocols, user roles	Safeguards data integrity and confidentiality
Data Storage	Storage solutions for fine-tuning data	Enables reliable storage and retrieval
Integration Tools	Tools to connect fine-tuning database with other systems	Facilitates seamless integration
Monitoring & Logging	Systems to track changes, access to data	Aids troubleshooting, optimization

In order to evaluate the performance of the GPT and LLM-based chatbot system proposed for technology credit evaluation of SMEs, a comprehensive evaluation was conducted using various indicators and verification techniques.

Various quantitative metrics were used to evaluate the accuracy, recall, and overall performance of the chatbot system. These metrics were chosen to measure both the model's ability to understand accurate user queries and the model's effectiveness, which provides relevant responses to SME technology credit ratings. In addition, a qualitative evaluation method was integrated to evaluate natural language responses generated in the chatbot system. The evaluator team reviewed the predefined questions raised in the chatbot system to evaluate appropriateness, naturalness, grammatical accuracy, and overall response quality.

To conduct evaluation experiments, prepare a test dataset consisting of real-world SME credit rating scenarios with varying levels of complexity. This dataset was classified into various technical areas, financial situations, management capabilities, market dynamics, and regulatory compliance requirements. The chatbot system was constructed with appropriate settings based on the training process using the fine-tuned GPT model specialized in technology credit evaluation of small and medium-sized enterprises. The model has been trained on a number of text data companies, consisting of industry reports, SME financial statements across various sectors, and regulatory documents related to small business credit rating practices.

The experimental results show the promising performance of the GPT-based chatbot system proposed in the technology credit evaluation of SMEs. As a result of quantitative analysis, the accuracy of accurately understanding user queries within the technical credit evaluation criteria defined by experts in the field was 92% compared to the previous evaluation. As a result of qualitative evaluation conducted by evaluators, the responses generated by the chatbot system were highly relevant (89% average relevance), natural language fluency (92% average fluency), and accurate understanding (91.2% understanding). In addition, through limited experiments on benchmark datasets commonly used in the field, the performance of the proposed model was compared with the existing state-of-the-art models used in SME credit ratings. This model demonstrates the rapidity of prioritizing technical reliability quickly and easily in terms of usability of this baseline model across various metrics such as reproducibility and speed of implementation.

## 6. Conclusion

This study examined the implementation of a GPT and LLM-based chatbot system for technical credit evaluation in SMEs. Through this approach, an automatic evaluation system was implemented relatively quickly compared to evaluation results through existing evaluation systems. The system demonstrated

situational judgment capabilities, adaptable data handling, and practicality, validated by application cases featuring real SMEs. However, the study acknowledges areas for improvement, including data diversification and enhanced user interfaces.

Limitations encompass limited data scope, system intricacy, and resource constraints. Constraints arise from data specificity, intricate integration of GPT and LLM, and resource limitations for SMEs.

Future research should focus on diverse data expansion, system optimization, cost-effective model development, and wider practical cases. Priorities involve generalizing the model with data from various industries, streamlining system complexity, and creating cost-effective & Quickly generative models tailored to SME characteristics.

## References

- [1] Ross Gruetzemacher, "The Power of Natural Language Processing, AI And Machine Learning," Harvard Business Review, <https://hbr.org/the-power-of-natural-language-processing>.
- [2] GPT-4 Technical Report, OpenAI
- [3] Bhattacharya, A., Biswas, S.K. & Mandal, A., "Credit risk evaluation: a comprehensive study," Multimed Tools Appl82, 18217–18267, 2023.  
DOI: <https://doi.org/10.1007/s11042-022-13952-3>
- [4] Hirk, R., Hornik, K. & Vana, L., "Multivariate ordinal regression models: an analysis of corporate credit ratings," Stat Methods Appl28, 507–539, 2019.  
DOI: <https://doi.org/10.1007/s10260-018-00437-7>
- [5] Automated Machine Learning Methods, Systems, Challenges, 2019.
- [6] G. Gao, H. Wang, and P. Gao, "Establishing a Credit Risk Evaluation System for SMEs Using the Soft Voting Fusion Model," Risks, Vol. 9, No. 11, 2021.  
DOI: <https://doi.org/10.3390/risks9110202>.
- [7] Roy, P.K., Shaw, K., "A multicriteria credit scoring model for SMEs using hybrid BWM and TOPSIS," Financ Innov7, 77, 2021.  
DOI: <https://doi.org/10.1186/s40854-021-00295-5>
- [8] E.I. Altman, M. Esentato, and G. Sabato, "Assessing the credit worthiness of Italian SMEs and mini-bond issuers," Glob Finance J, 2017:100450,  
DOI: <https://doi.org/10.1016/j.gfj.2018.09.003>.
- [9] Linardatos P, Papastefanopoulos V, Kotsiantis S., "Explainable AI: A Review of Machine Learning Interpretability Methods," Entropy. Vol. 23, No.1, p. 18, 2021.  
DOI: <https://doi.org/10.3390/e23010018>
- [10] J.A. Bargas-Avila and K. Hornbæk, "Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience," in Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011), pp. 2689–2698, ACM, 2011.
- [11] V. Dey, "ChatGPT and LLM-based chatbots set to improve customer experience," VentureBeat, <https://venturebeat.com/ai/chatgpt-and-llm-based-chatbots-set-to-improve-customer-experience/>.
- [12] AIMultiple, "How to Build a Chatbot: Components & Architecture in 2023 - AIMultiple," <https://research.aimultiple.com/chatbot-architecture>.