AI-Powered Assessment of Resistance to Change in the Context of Digital Transformation

Bachira Abou El Karam, Tarik Fissaa, Rabia Marghoubi
Innovative Research on Software Systems and Data, INPT, Rabat, Morocco

***Abstract*—Digital transformation is a key driver of business evolution, but it comes with significant challenges, particularly employee resistance to change. This resistance can manifest in various forms, ranging from explicit opposition to more subtle hesitation toward new practices. Its underlying causes are diverse, including fear of the unknown, loss of control, and dissatisfaction with perceived transformations. Understanding employee perceptions is, therefore, crucial to adapting digital initiatives and ensuring successful adoption. However, existing methods for assessing resistance, which rely on closed-ended questionnaires and binary classifications, have limitations. They restrict the expression of opinions and fail to provide a nuanced segmentation of employees' stances toward change. In this context, this study proposes an innovative and automated methodology that combines specialized zero-shot LLMs and prompt engineering techniques to analyze resistance to change. It is based on the allies strategy, a concept derived from sociodynamic theory and widely applied in change management, which seeks to more precisely differentiate employee attitudes based on their level of synergy or antagonism toward a new project or transformation initiative. To evaluate the effectiveness of the proposed approach, an experiment was conducted on an annotated dataset comprising a hundred employee responses. Two prompt engineering strategies were explored and applied to six zero-shot models to assess their ability to accurately classify expressed attitudes. The findings underscored, on one side, the significance of prompt structuring in enhancing classification efficacy and, on the other side, the preeminence of DeBERTa-v3-large-zeroshot, which demonstrated itself as the most exemplary model, even exceeding GPT-4, one of the most sophisticated and cutting-edge language models currently accessible.**

Keywords*—*Resistance to change; digital transformation; zero-shot LLMs; prompt engineering; allies strategy

I. Introduction

Resistance to change is a widely studied phenomenon in management and organizational psychology. It is generally defined as a cognitive, emotional, or behavioral reaction to a modification perceived as a threat to stability, habits, or the status quo [1] [2] . M. Asem [3] describes it as opposition to new organizational practices, which may be driven by personal interests, a lack of understanding, low tolerance for uncertainty, or a different assessment of the proposed change.

In the sphere of digital transformation, resistance to change takes on a specialized form. It is often linked to technological upheavals that redefine processes, roles, and interactions within organizations. Among the primary determinants contributing to this resistance encompass job insecurity, as employees may possess uncertainties regarding prospective layoffs or modifications in their roles; inadequate technological aptitude, which can elicit apprehension and negative perceptions towards digital technologies; and issues related to identity and social relationships, with particular technological innovations regarded as threatening one's social standing or reputation within the organization [4][5][6][7][8][9]. These factors demonstrate that resistance to change in a digital context is not solely based on irrational fears, but on legitimate concerns that must be considered in change management strategies [10].

Resistance to change has often been conceptualized from two main perspectives in literature: a systemic approach, which views resistance as an organizational phenomenon influenced by corporate culture, processes, and structure, and an individual-centered approach, which examines how employees perceive and react to change [8][11]. However, these two approaches are closely linked. In [12] and [13], the authors highlight that organizational-level resistance often leads to individual resistance, as employees react to transformations perceived as being imposed by the existing system.

Thus, while research on systemic resistance provides a critical framework for understanding structural obstacles to digital transformation [5][6][7][9][14][15], it is at the individual level that resistance takes a concrete form and directly influences the success or failure of a change initiative. Therefore, it is essential to focus on methods for measuring this resistance on the individual scale. In this context, several measurement tools have been developed to assess employee resistance to change [16][17][18][19][20]. These instruments, often in the form of closed questionnaires or scales, have the advantage of enabling quantitative and comparative assessment of resistance to change. However, their drawback is that they unfortunately limit the richness of responses and do not fully capture the diversity of perceptions and emotions employees have towards change. They tend to reduce the range of responses to a simple dichotomy between resistors and non-resistors, leading to an imprecise or biased classification that does not faithfully reflect the complexity of reactions to change. This reduction of nuances hinders the identification of intermediate profiles and complicates the development of strategies tailored to each stakeholder in the change process.

In response to the limitations of traditional approaches, we propose an alternative method based on the use of the allies strategy and zero-shot language models. Instead of relying solely on closed-ended questionnaires or rigid classifications, our approach focuses on analyzing employees' open-ended responses, harnessing the power of zero-shot LLMs combined with prompting techniques to perform a deep natural language analysis. At the same time, the allies strategy helps to better identify and segment employees based on their potential role in the transformation: those actively promote change and those who may hinder it. This segmentation will enable communications and change management actions to be targeted effectively, making it easier to overcome resistance and build commitment to the digital initiatives arising from digital transformation.

The following sections in the study are arranged as: Section II presents related work. Section III provides essential background information on zero-shot LLMs, prompt engineering, the allies strategy, and foundational concepts crucial for understanding the context of our study. Section IV presents the methodology employed in measuring resistance to change using the allies strategy, including the design and implementation of prompt engineering strategies and the selection of zero-shot LLMs. The experimental setup is elaborated in Section 5V, while Section VI elucidates the results obtained and their subsequent interpretation. In closing, Section VII recaps the research and outlines potential routes for further examination.

II. Related Work

In the literature, the assessment of resistance to change in the context of digital transformation has been approached through two primary lenses. From one angle, a holistic viewpoint underscores the importance of pinpointing organizational aspects that contribute to resistance, their effects on business success, and the techniques for identifying and reducing that resistance. Conversely, a person-focused perspective delves into how employees react to digital shifts by exploring their viewpoints, hurdles, and incentives. However, a common limitation of many studies lies in their tendency to adopt a generalized view that does not always account for the richness of expressions and the freedom employees have to articulate their own perceptions of change, nor the variability in their stance toward it.

Several studies have taken a systemic approach in analyzing resistance to change in the context of digital transformation without analyzing in depth how it manifests within employees, for instance, both studies [5] and [6] highlight common sources of resistance such as fear of job loss, lack of digital skills, perceived work overload, and concerns about professional identity. However, these studies remain at a macro level, failing to segment employees based on their specific resistance attitudes or to propose a classification of resistance profiles that could support more tailored change management strategies. In a related vein, Tetiana Kuzhda [14] proposes a diagnostic model of resistance to change based on expert surveys to identify and prioritize the major causes of resistance and then suggest strategies for reducing it. Still, these methods are rather standard and overlook the personal differences among workers. Finally, Samia Hattab [15] investigates the issue of resistance to change in the public sector by analyzing a survey of administrative staff to determine the primary overall causes of this resistance during digital transformation. However, this survey highlights general trends and fails to categorize employees according to their degree of resistance, which limits the possibility of conceiving targeted change management strategies.

Consequently, it is necessary to concentrate on methodologies for quantifying resistance at the individual scale. From this viewpoint, numerous approaches have been formulated to evaluate employees' resistance to organizational change. One of the most widely used tools is the Resistance to Change (RTC) scale developed by Oreg [16], which measures individual resistance across four dimensions: routine-seeking, emotional attachment, reaction to stress, and reluctance to lose control. This approach relies on standardized questionnaires that quantify perceived resistance. In the particular context of digital transformation, numerous research contributions have endeavored to enhance the assessment of resistance to change by considering the unique characteristics of digital contexts. For example, Sherif Hamdi [17] examines employee resistance within the National Social Insurance Organization in Egypt, and analyzes its role as a mediating variable between training and performance. However, this study relies on closed-ended questionnaires (agree or disagree), which limits a nuanced understanding of employees' individual barriers and motivations regarding change.

In a different approach, Stam et al.[18] uses Membership Category Analysis (MCA) to analyze employee discourse on the introduction of digital technologies. This qualitative approach highlights the discursive categories used to express their perceptions of change. However, although this method offers a detailed analysis of social representations, it relies entirely on human interpretation, which limits its large-scale application and does not allow for precise segmentation of employees according to their degree of resistance.

Another attempt at classification is proposed in our previous work [19]. This study relies on generalized hesitant fuzzy sets and Formal Concept Analysis to classify employees based on their level of acceptance of change, incorporating uncertainty in response formulation. Nevertheless, although this methodology affords increased liberties of discourse in comparison to closed-ended inquiries, it does not entirely exploit the depth of open-ended replies and does not facilitate a varied typology of employee profiles.

Finally, the study [20] explores another variant of fuzzy sets by introducing an intuitionistic analysis to capture resistance to change in digital contexts. This approach enhances the modeling of hesitations and uncertainties in employees' responses but remains dependent on a rigid mathematical framework that does not adapt to semantic variations and the subtleties of natural language.

The analysis of existing work reveals several limitations. Some studies translate the analysis and assessment of resistance to change in the context of digital transformation by identifying the causes of resistance and thus propose general recommendations applied to all employees. Other studies that have sought to analyze employee's perceptions of the introduction of new digital transformation projects have generally adopted a binary classification of employees as resistors and non-resistors, neglecting intermediate positions such as those who are “torn”, “conciliatory” or “indecisive”, which limits the identification of more nuanced categories and hinders the implementation of targeted strategies adapted to each change actor. What's more, the methodologies employed using closed-ended questionnaires do not fully capture the complexity of individual attitudes to transformation, which can lead to imprecise or biased classification.

It is within this perspective that our contribution is situated. We propose adopting the allies strategy, derived from the sociodynamic approach developed by Jean-Christian Fauvet in the 1970s [21], to establish a more granular classification of employees in relation to change, based on open-ended responses. Indeed, and unlike traditional approaches that often rely on binary segmentation and closed questions, preventing a more nuanced understanding of employees' positions towards digital transformation, our approach enables us to identify different stakeholder profiles while promoting employees' freedom of expression thanks to open-ended responses. This analysis will enable the development of a customized change management plan and help organizations to target their communication, training and support actions more effectively, thereby reinforcing buy-in to the transformation. To automate this approach, Large Language Models (LLMs) specialized in zero-shot classification, combined with prompting techniques, were used to analyze and process open-ended responses.

III. Background

This section presents fundamental background information on the key topics that form the basis of our study, namely the sociodynamic approach, the allies strategy, Zero-shot LLMs and prompt engineering.

A. The Sociodynamic Approach

Developed by Jean-Christian Fauvet in the 1970s, the sociodynamic approach is a strategic method for analyzing human dynamics and individuals' attitudes towards change based on the idea that any transformation generates power relationships and interactions between different actors with varied interests [21]. Unlike traditional change management approaches, which prioritize top-down and prescriptive methodologies, this method adopts an interactionist and adaptive perspective [22]. It considers that the success of a transformation project depends not merely on technical and organizational aspects but also on relational dynamics and stakeholder engagement strategies. This dynamic is based on theoretical foundations such as the theory of commitment [23] and the theory of influence networks [24], which explain how social interactions and gradual involvement alter attitudes towards change.

As Paul Walley has highlighted [25], applying a sociodynamic perspective to stakeholder management enhances project outcomes by enabling a more nuanced understanding of stakeholder behaviors and by effectively addressing resistance to change. Indeed, the sociodynamic approach allows for the recognition of complex, and sometimes ambivalent, employee responses, where individuals may express both support and hesitation simultaneously, offering a richer and more realistic assessment of attitudes during organizational transformations.

B. The Allies Strategy

The allies strategy model is an operational application of the sociodynamic approach, specifically designed to classify actors based on their attitude towards a transformation project into allies, neutrals, undecided, and opponents, thus enabling a differentiated strategy depending on their posture. This approach is structured around two axes: the synergy axis, which aims to mobilize change supporters to positively influence the undecided, and the antagonism axis, which involves identifying and mitigating the impact of opponents [25] [26].

The central tool of the model, known as the actor mapping or partner map [25] [26] [27] [28], allows for a more detailed analysis of individual resistance, facilitating the implementation of tailored strategies to foster engagement and support the transition. It aids in illustrating the diverse perspectives of individuals or groups upon their degree of synergy (support) and antagonism (resistance) regarding the initiative. This mapping is often represented as a matrix, as shown in Fig. 1, consisting of two main axes: The first axis, which measures the degree of support or commitment an actor has towards the project, is subdivided into four main categories:

- Minimalist: Individuals with little interest or commitment to the project, showing minimal involvement.
- Interested: Individuals who show a degree of interest but require more evidence or benefits before fully committing.
- Cooperative: Individuals who are actively engaged in the project, offering ideas and solutions.
- Committed: Those who reflect a strong devotion to the project, are motivated, and are inclined to assume an important position in its operationalization.

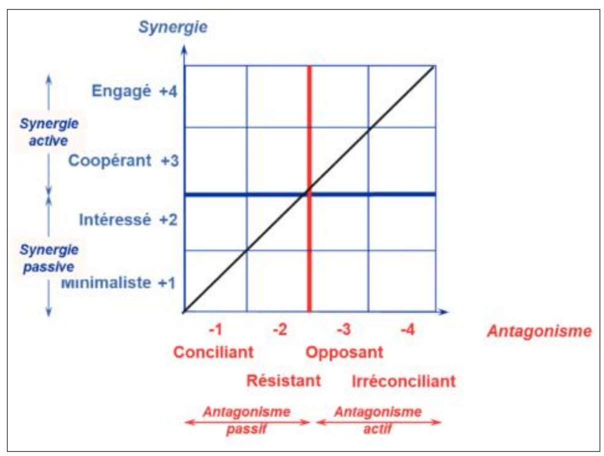

Fig. 1. Partner map [26].

The second axis, which evaluates the level of opposition or resistance an actor has towards the project, is also subdivided into four main categories:

- Conciliatory: Individuals with reservations about the change but open to dialogue and compromise.
- Resistant: Individuals who show stronger opposition due to specific concerns or divergent interests.

- Opponent: Individuals who actively oppose the project, potentially seeking to block it due to personal fears or general resistance to change.
- Irreconcilable: Individuals whose opposition is deep and often irrevocable, making constructive involvement difficult.

The intersections between the categories of synergy and antagonism give rise to seven more nuanced and meaningful categories, reflecting the complexity of organizational and human dynamics during change processes.

One of the main added values of this model resides in its capacity to reveal concealed or passive stakeholders, individuals who may not be readily apparent or vocal during the change process, yet whose perspectives and influence can profoundly affect project outcomes. By rendering these latent dynamics discernible, the allies strategy prompts project managers to expand their stakeholder analysis beyond the most conspicuous participants, thereby facilitating more tailored engagement strategies and mitigating the risk of neglecting silent sources of resistance or support.

C. Zero-Shot LLMs

A Large Language Model (LLM) represents an advanced manifestation of artificial intelligence, systematically engineered to analyze and generate natural language akin to human discourse. Based on sophisticated deep learning techniques, particularly on an advanced neural network architecture known as Transformers [29], these models incorporate an attention mechanism that discerns complex interrelationships between textual components, regardless of their position in the sequence. The effectiveness of this attention mechanism gives LLMs the ability to capture the nuances of language, syntax, and contextual relevance. Owing to their thorough preliminary training on a vast number of datasets [30], LLMs are exceptionally effective in numerous applications, which include conversational agents, automated content production, translation, and what is termed 'zero-shot' classification, defined as a technique that permits a model to label text without prior familiarity with examples of the pertinent category [31] [32]. LLMs fine-tuned to this specific task are called zero-shot LLMs, and they use this technique to correlate textual data with classifications without prior exposure to such data. They rely on natural language inference (NLI) models [33], which investigate the semantic relationships between an input text and a particular category through a logical semantic process that assesses whether a certain hypothesis (the category) is entailed (implication), neutral (neutrality), or contradicted (contradiction) by a provided text [34]. The performance of these models is chiefly derived from their foundational training on vast, high-standard natural language inference (NLI) datasets [32], incorporating SNLI [34], MNLI [35], ANLI [36], and FEVER [37]. These datasets contain extensive collections of annotated sentence pairs, assessing whether one sentence entails, is neutral to, or contradicts another. They have enabled models to develop a more refined understanding of semantic relationships and enhance their inference capabilities.

Within the scope of our research, the adoption of zero-shot LLMs is notably supported by the open-ended and unstructured responses from the feedback gathered from employees related to their views on organizational change. Zero-shot models enable the immediate classification of such qualitative responses without the need for prior task-specific fine-tuning. Furthermore, their capacity to perform inference-based categorization aligns well with the conceptual framework of the allies' strategy, which requires distinguishing nuanced and context-dependent postures ranging from strong support to strong opposition. This makes them highly appropriate for analyzing resistance to change in dynamic and resource-constrained environments.

D. Prompt Engineering

Prompt engineering constitutes a technique for optimizing the utilization of large language models (LLMs) by steering their outputs via precise directives, all while preserving the integrity of the model itself. This technique enables the optimization of LLMs' performance in various tasks, such as text classification, machine translation, or generating tailored responses, simply by adjusting the phrasing of the queries [38]. What renders prompt engineering exceptionally potent is its capacity to adapt pre-trained models for novel tasks without necessitating the retraining of the model parameters. As articulated in [39], this methodology significantly augments the capacity of models to generalize proficiently to novel tasks, especially via mechanisms such as zero-shot prompting and few-shot prompting. The effectiveness of this technique depends on the precision of prompt formulation, as even small adjustments can lead to significant variations in the results generated [40]. Several prompting strategies have been developed, including, for example:

- Zero-Shot Prompting: The model is given an instruction without any example responses. It relies on the LLM's ability to generalize from existing knowledge [38].
- Few-Shot Prompting: This strategy includes several representative instances within the prompt to improve the model's grasp of the task at hand. It enhances model accuracy by providing more context [38].
- Chain-of-Thought Prompting (CoT): As delineated by J. Wei et al [39], this methodology invites the model's elucidation of its reasoning process via intermediary steps prior to delivering a conclusive response. It’s particularly useful for tasks requiring complex logical reasoning, such as math problems or advanced inference.
- Role-Playing Prompting: The model is assigned a specific role, such as an expert in a given domain, to improve the relevance and coherence of responses in specialized contexts [40].
- AI-Knowledge Prompting: As articulated by J. Liu et al. [41], this methodology entails soliciting the model to produce knowledge that is specific to a given task, which is subsequently integrated into the prompt to enhance both the precision and reliability of the generated predictions.

IV. Methodology

Our methodology relies on the use of zero-shot LLMs combined with prompting techniques to analyze and classify open-ended responses in the context of measuring resistance to change, based on the allies strategy. The choice of these models is justified by their ability to infer logical relationships between texts and recognize classification patterns without requiring supervised learning on a specifically annotated dataset [33]. However, their performance is highly dependent on the formulation of hypotheses and the provided context [42] [43]. In order to enhance classification accuracy, we have incorporated a prompt engineering methodology aimed at optimizing the models' comprehension [44].

In the following, we give a brief description of the zero-shot LLMs selected for our study, and we also describe the prompt formulations used, followed by detailing the workflow of our proposed approach to measuring resistance to change relying on allies strategy.

A. Models

In the context of our research, we chose a selection of LLMs specialized in zero-shot classification and recommended by the Hugging Face platform [45] due to their performance in natural language inference (NLI), namely:

1) Facebook/bart-large-mnli. The BART model, known as Bidirectional and Auto-Regressive Transformers, represents an encoder-decoder transformer that integrates the benefits of bidirectional encoder models like BERT alongside auto-regressive decoder models such as GPT [46]. Its variant, facebook/bart-large-mnli, has been pretrained on the MultiNLI dataset, enabling it to perform excellently in Natural Language Inference (NLI) tasks. Its architecture makes it particularly well-suited for zero-shot classification, where it evaluates the probability that a given text corresponds to a predefined hypothesis, a mechanism directly leveraged in our study to assign responses to categories of resistance to change.

2) MoritzLaurer/deberta-v3-large-zeroshot-v2.0 and MoritzLaurer/deberta-v3-base-zeroshot-v2.0mnli. eBERTa models (Decoding-enhanced BERT with Disentangled Attention) are an improvement over BERT and RoBERTa, introducing a disentangled representation mechanism that more effectively distinguishes words based on their position and context [47]. The variants deberta-v3-large-zeroshot-v2.0 and deberta-v3-base-zeroshot-v2.0, developed by Moritz Laurer, are specifically optimized for zero-shot classification. They benefit from enhanced masking and more efficient pre-training, thereby improving their ability to generalize to new tasks without explicit supervision.

3) MoritzLaurer/roberta-large-zeroshot-v2.0. oBERTa (Robustly Optimized BERT Pretraining Approach) is an improved variant of BERT that relies on longer pretraining and better hyperparameter tuning [48]. The roberta-large-zeroshot-v2.0 version, fine-tuned for zero-shot classification, leverages these optimizations to better handle the variability of natural language, an essential advantage in our task, where employee responses can vary significantly in wording and tone.

4) MoritzLaurer/bge-m3-zeroshot-v2.0. The bge-m3-zeroshot-v2.0 model is part of the models optimized for semantic similarity tasks and zero-shot classification. It relies on a combination of advanced semantic representation techniques and specific optimizations to improve robustness when dealing with unstructured texts and varied formulations [4].

5) OpenAI/GPT-4. GPT-4 is a multimodal model and one of the most advanced developed by OpenAI, outperforming its predecessors in terms of contextual understanding, logical reasoning and text generation [49]. Unlike specialized models such as BART or DeBERTa, GPT-4 is a generalist model, but it excels in zero-shot classification tasks thanks to its ability to analyze global context and interpret complex formulations. Its integration into our study enables us to assess the extent to which an ultra-polyvalent model can compete with specialized models on a specific task, such as the analysis of resistance to change.

B. Prompt Engineering

To enhance the zero-shot classification of employees' responses while relying on the allies strategy, we have developed and tested two types of prompts, namely a minimalist prompt and an explicit prompt:

1) A minimalist prompt. This prompt is limited to essential information, represented solely by numeric labels and class names corresponding to the synergy and antagonism axes as defined in the allies strategy. In this scenario, two sets of classes were defined according to this strategy:

- Synergy (synergy class): "1 Minimal", "2 Interested", "3 Cooperative", "4 Committed"
- Antagonism (antagonism class): "1 Conciliatory", "2 Resistant", "3 Opponent", "4 Irreconcilable"

This type of prompt aims to evaluate the zero-shot models' ability to generalize from minimal class descriptions, without providing explicit context about their meaning. In this case, we rely entirely on LLM’s internal representation and its pre-existing knowledge gained from vast textual datasets to generalize new classification tasks. This approach is aligned with the Zero-Shot Prompting strategy.

2) An explicit prompt. This approach is based on AI-Knowledge Prompting [41], where each class (synergy and antagonism classes) is enriched with an explicit definition of the behaviors associated with each level of synergy or antagonism. These descriptions provide the model with a more precise contextual framework, thereby facilitating the distinction between different classes. To design these AI-Knowledge Prompts, we used GPT-4 to generate detailed class descriptions. Specifically, we formulated an initial zero-shot prompt containing only the classification labels, and then we leveraged GPT-4 to generate definitions specific to each category, based on the principles of the allies strategy. These descriptions were subsequently integrated into the final explicit prompt to provide a richer contextual framework.

- Synergy (synergy class):
  - "Follows assigned directives without active initiatives",
  - "Follows assigned directives and expresses verbal interest in the project, but without active initiatives",
  - "Takes active initiatives",
  - "Takes active initiatives with a high level of responsibility".
- Antagonism (antagonism class):
  - "Supports the project’s decisions and approaches",
  - "Opposes the project or certain decisions/approaches but is open to discussion",
  - "Opposes the project or certain decisions/approaches and is less flexible in discussions",
  - "Opposes the project or certain decisions/approaches and is inflexible in discussions".

This type of prompt aims to examine how the detail of class descriptions affects the efficiency of zero-shot models.

C. Proposed Approach for Measuring Resistance to Change

In this section, we provide an overview of the methodology employed in our study, illustrated in Fig. 2, detailing the steps taken to investigate the effectiveness of zero-shot LLMs combined with prompting techniques in measuring resistance to change using the allies strategy. We adopted a multi-task classification approach, with each task focusing on a different aspect of employees' dynamics regarding digital transformation projects:

- Classification based on Synergy Level: The objective of this task is to categorize employees by their degree of engagement and collaboration within the project. The labels range from "Minimalist" to "Committed". This classification helps identify the most motivated team members and those who may require additional motivation or incentives.
- Classification based on Antagonism Level: focuses on analyzing the degree of resistance or opposition exhibited by employees regarding project initiatives. The classifications range from 'Conciliatory' to 'Irreconcilable', thus enabling the recognition of resistance levels that could affect group dynamics and hinder project progress.
- Final classification: This task synthesizes results derived from the initial two classifications to construct a holistic appraisal of each employee. It amalgamates both synergistic and antagonistic dimensions to yield a comprehensive evaluation of employees' attitudes towards digital transformation initiatives.

Based on a dataset containing open-ended employee responses, previously annotated with synergy and antagonism classes, we proceed as follows:

The first two tasks, namely classification by Synergy Level and Classification by Antagonism Level, referred to both simultaneously in Fig. 2 as task 1 and task 2, follow the same process, detailed below:

1) Preprocessing of open-ended responses. Each open-ended response provided by an employee is standardized to ensure better understanding by the language models. This step includes text cleaning (removal of special characters, unnecessary stopwords if necessary) and conversion into a standardized format suitable for input into zero-shot classification models.

2) Prompt engineering. Two variants of prompts are used to guide the zero-shot classification models:

a) Minimalist prompt. Contains only the classification labels.

b) Explicit prompt. Adds detailed descriptions of the classes to improve the model's interpretation.

3) Zero-shot classification with LLMs. Each open-ended response is submitted to a large language model (LLM) specialized in zero-shot classification, accompanied by the selected prompt (either minimalist or explicit). The model reformulates the task as a natural language inference (NLI), where the open-ended response is considered the "premise" and each possible class is treated as a "hypothesis". It then evaluates the probability that the response implies each hypothesis, assigning a confidence score to each class. Finally, the class with the highest probability is selected as the final label assigned to the open-ended response.

4) Performance evaluation of classification. The classification performance is evaluated by comparing the predicted classes to the actual classes annotated in the dataset. The metrics calculated are:

- Confusion Matrix
- F1-score
- Recall
- Accuracy
- ROC-AUC Curve

Regarding the third classification task (task 3 in Fig. 2), which aims to assign the final class of resistance to change, the process is as follows: Once the synergy and antagonism classes have been determined according to the steps of the first two tasks, they are combined based on a pre-established matrix derived from the allies’ map. This combination allows for the assignment of a final class representing the overall level of resistance to change.

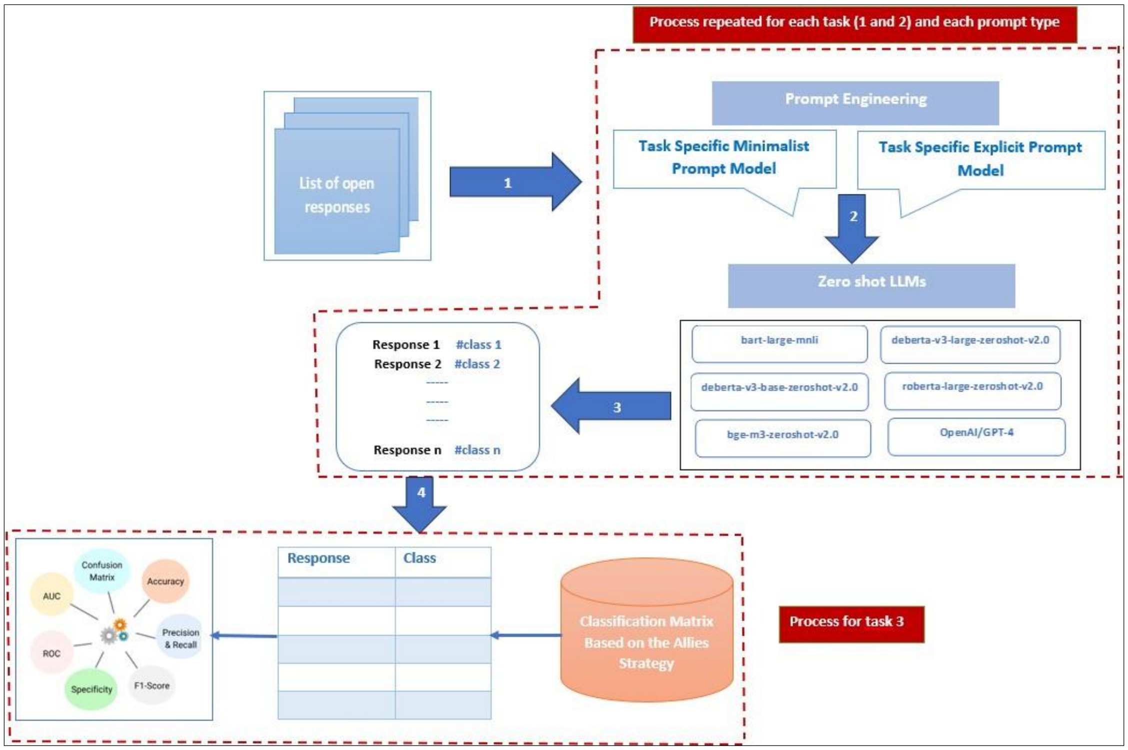

Fig. 2. Overview of the approach for measuring resistance to change.

As with the previous tasks, the predicted final class is compared to the actual class annotated in the dataset. Performance evaluation is carried out using the same metrics as the previous tasks.

V. Experiments

To validate the proposed methodology, we executed a thorough array of experiments intended to measure the effectiveness of the selected zero-shot Large Language Models (LLMs) in precisely classifying employee opinions regarding the context of digital transformation, derived from their open-ended text replies. This section elucidates the experimental framework, encompassing the dataset utilized, the evaluative methodology employed to benchmark the models, and the metrics implemented to evaluate their efficacy. The design of these experiments aimed to investigate both the efficacy of the prompting strategies and the pertinence of the allies strategy framework in effectively capturing the intricate expressions of resistance to change.

A. Dataset

As part of our study aimed at classifying employees based on their synergy and antagonism towards a digital transformation project, we built a database by collecting approximately one hundred open-ended responses directly from a company that recently introduced digital initiatives. This data collection allowed us to gather a hundred employee testimonies expressing their perceptions, concerns, and levels of adherence to these changes. Each open-ended response was then carefully annotated by change management experts in accordance with the classes defined in the allies strategy. Specifically, three labels were assigned to each response:

- A synergy level, indicating the degree of alignment and positive engagement of the employee towards the digital initiative.
- An antagonism level, reflecting the resistance expressed, whether passive or active, towards the change.
- A global resistance to change level, resulting from the combination of the first two classifications as defined in the Allies Strategy Matrix.

These manual annotations are an essential benchmark for evaluating the performance of the models used, enabling us to measure their ability to reproduce expert judgments in the analysis of resistance to change.

The distribution of employee responses by level of synergy and antagonism is described in Fig. 3:

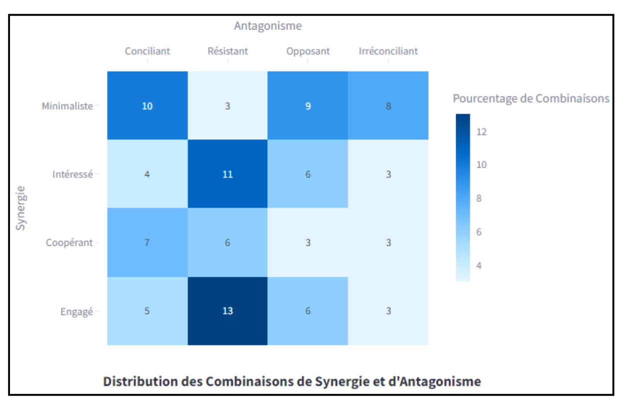

Fig. 3. Distribution of employee responses by level.

B. Evaluation Process

Our evaluation process, as described in the methodological section, will be applied to six models specialized in zero-shot classification, using two distinct prompt variants for each tested model: a minimalist prompt containing only the classification labels and an explicit prompt providing detailed descriptions of the classes to enhance the model's interpretation. In other words, each model is required to classify the open responses of the dataset according to the level of synergy corresponding to task 1 and then according to the level of antagonism corresponding to task 2, in line with the workflow described in the methodology section. These classifications, predicted by the models, will be generated first using the minimalist prompt and then the explicit prompt. Finally, once the synergy and antagonism classes have been determined according to the steps of the first two tasks. Task 3 is started by combining the results of the first two tasks based on a pre-established matrix derived from the allies strategy. This combination makes it possible to assign a final class representing the overall level of resistance to change for each open response.

C. Evaluation Metrics

To validate the performance of the models, we use a method structured in two main steps:

1) Evaluation by task. Evaluation by task consists of comparing the classifications predicted by the model with the actual classifications annotated in the dataset for each of the tasks and prompts defined in our classification workflow. For this, we use confusion matrices to visualize the distribution of errors made by the model, and we also use standard classification metrics, notably:

a) Precision: This metric is crucial for measuring the reliability of the model's positive classifications. In our context, a high precision score indicates that when the model identifies a specific stance toward change, it is unlikely to be incorrect. This is particularly important to avoid misclassifications that could distort the analysis of resistance and support for digital initiatives.

b) Recall: This metric evaluates the model's proficiency in detecting all affirmative instances present within the dataset. A high recall score is essential in our study to ensure that minority, but potentially critical, stances are not overlooked. This helps to provide a more comprehensive understanding of the dynamics of resistance and synergy among employees.

c) F1-score: By combining precision and recall, the F1-score balances these two aspects, particularly in cases where class distribution is imbalanced. In our analysis, this metric is particularly relevant, as some stance categories may be underrepresented, making it necessary to use a measure that penalizes both false positives and false negatives.

d) Accuracy: This metric assesses the ratio of correctly predicted classifications relative to the total predictions generated by the model. While it offers a general measure of reliability, it may not fully reflect performance in cases of class imbalance.

2) Comparative evaluation of models. Beyond analyzing individual classification tasks, we conduct a comprehensive comparison of the models using two key performance indicators:

a) The F1 score for the final classification task: This indicator summarizes the model's overall ability to balance precision and recall in assigning final classes.

b) Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC): Through the examination of the equilibrium between true positives and false positives, the AUC score serves to ascertain the model's ability to consistently differentiate among various levels of resistance, which is imperative within our framework, where a nuanced classification is essential for informed decision-making in developing a change management plan tailored to each category.

By combining these different metrics, we are able to identify the best-performing model best suited to our problem, thus guaranteeing accurate and exploitable results for the analysis of resistance to change.

VI. Results and Discussion

In this section, we evaluate the ability of the models to correctly classify responses according to the defined tasks (synergy, antagonism). For each task and each type of prompt, we have calculated the standard classification metrics: precision, recall, F1 score and accuracy, along with confusion matrices for fine-grained error analysis.

The performance of the models using both types of prompts is presented in Table I. Metrics are calculated using the "macro" method, which means that we calculate precision, recall, accuracy and F1 score for each class independently, then average them. This approach better reflects the overall performance of the model across all classes.

The preliminary analysis of the results highlights the following points:

- Minimalist prompt: Overall, the model's performance is moderate to low on both tasks when using a minimalist prompt (see Table I). This trend suggests that overly basic formulations do not provide enough context or information to the models, limiting their ability to correctly distinguish between classes.
- Explicit prompt: Adding more detailed descriptions significantly improves performance, particularly in terms of the F1 score, as illustrated in Fig. 4 and Fig. 5 for nearly all models. This improvement is particularly notable for the DeBERTa-v3-large-zeroshot model, which achieves the highest F1 scores on both tasks, demonstrating the importance of a well-structured prompt to optimize classification.

The results also show that the explicit prompt yields the best overall performance, with a notable increase in both precision and recall as shown in Table I.

TABLE I. PERFORMANCE METRICS

Model	Prompt	Task 1: Synergy Classification				Task 2: Antagonism Classification			
		Precision	*Recall*	*F1 Score*	*Accuracy*	*Precision*	*Recall*	*F1 Score*	*Accuracy*
OpenAI/GPT4	***Minimalist***	**0.52**	**0.43**	**0.40**	**0.46**	**0.64**	**0.50**	**0.45**	**0.56**
	Explicit	**0.70**	**0.62**	**0.60**	**0.65**	**0.72**	**0.60**	**0.60**	**0.65**
Facebook/bart-large-mnli	***Minimalist***	**0.35**	**0.32**	**0.22**	**0.32**	**0.26**	**0.43**	**0.31**	**0.43**
	Explicit	**0.53**	**0.46**	**0.44**	**0.47**	**0.60**	**0.49**	**0.42**	**0.56**
MoritzLaurer/deberta-v3-large-zeroshot-v2.0	***Minimalist***	**0.46**	**0.46**	**0.40**	**0.50**	**0.41**	**0.45**	**0.38**	**0.44**
	Explicit	**0.76**	**0.76**	**0.76**	**0.77**	**0.84**	**0.79**	**0.80**	**0.81**
MoritzLaurer/roberta-large-zeroshot-v2.0	***Minimalist***	**0.45**	**0.44**	**0.42**	**0.47**	**0.56**	**0.45**	**0.45**	**0.47**
	Explicit	**0.67**	**0.64**	**0.64**	**0.66**	**0.66**	**0.67**	**0.65**	**0.69**
MoritzLaurer/deberta-v3-base-zeroshot-v2.0	***Minimalist***	**0.42**	**0.33**	**0.28**	**0.31**	**0.55**	**0.47**	**0.41**	**0.45**
	Explicit	**0.51**	**0.47**	**0.44**	**0.50**	**0.66**	**0.54**	**0.53**	**0.53**
MoritzLaurer/bge-m3-zeroshot-v2.0	***Minimalist***	**0.38**	**0.37**	**0.29**	**0.38**	**0.39**	**0.35**	**0.35**	**0.33**
	Explicit	**0.41**	**0.47**	**0.41**	**0.48**	**0.59**	**0.47**	**0.45**	**0.48**

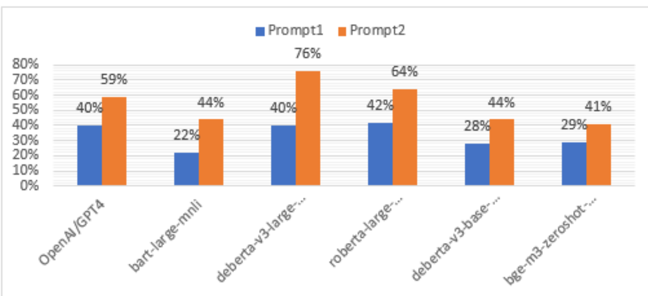

Fig. 4. F1 Score Results for Task 1 according to both prompts.

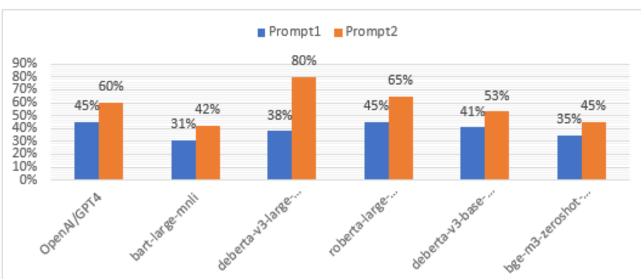

Fig. 5. F1 Score Results for Task 2 according to both prompts.

Based on the results shown in Fig. 4 and Fig. 5, it is clear that the explicit prompt (referred to in Fig. 4 and 5 as Prompt2) provides better performance than the minimalist prompt (referred to in Fig. 4 and 5 as Prompt1). To further analyze this, we will focus solely on the explicit prompt and examine the errors made by the models on tasks 1 and 2 using confusion matrices. These matrices allow us to visualize the correspondence between the model predictions and the actual classes, highlighting classification errors. The goal is to identify the models that achieve the best results and those that exhibit recurring errors.

To do this, we will display and analyze the confusion matrices of the different models on these two tasks in order to visualize the distribution of errors and better understand the respective performance of each model.

For task 1 related to classification according to the synergy axis, we selected the confusion matrices of the six models analyzed, as shown in Fig. 6.

The analysis of the confusion matrices reveals several types of recurring errors. Firstly, some models tend to confuse intermediate classes, particularly "Cooperative" and "Interested", suggesting difficulty in capturing the nuances between these categories. For instance, BART-large-mnli and RoBERTa-large-zeroshot frequently confuse these two classes, while DeBERTa-v3-base-zeroshot shows notable errors by incorrectly assigning "Cooperative" responses to "Committed".

Additionally, the BGE-m3-zeroshot model appears to struggle with correctly identifying the "Minimalist" class, redistributing its predictions toward other categories. In contrast, DeBERTa-v3-large-zeroshot and GPT-4 stand out with better classification performance, with DeBERTa-v3-large-zeroshot displaying slightly higher accuracy. However, errors persist in the intermediate classes, indicating a challenge in clearly differentiating these categories.

For task 2, related to classification according to the axis of antagonism, we adopt the same approach by generating the confusion matrices of the six models, as illustrated in Fig. 7.

Based on the analysis of the confusion matrices in Fig. 7, it is observed that DeBERTa-v3-large and GPT-4 stand out for their ability to accurately classify the extreme classes, with high precision for the categories "Conciliatory" and "Resistant". DeBERTa-v3-large demonstrates a more stable classification across all intermediate classes compared to the other models, while GPT-4 shows some confusion in the categories "Opponent" and "Irreconcilable". BART-large-mnli and RoBERTa-large-zeroshot also manage to correctly identify the majority classes but exhibit more pronounced confusion in differentiating the intermediate categories. DeBERTa-base and BGE-m3-zeroshot, on the other hand, seem to struggle more with distinguishing the "Opponent" and "Irreconcilable" classes, with a more scattered distribution of errors.

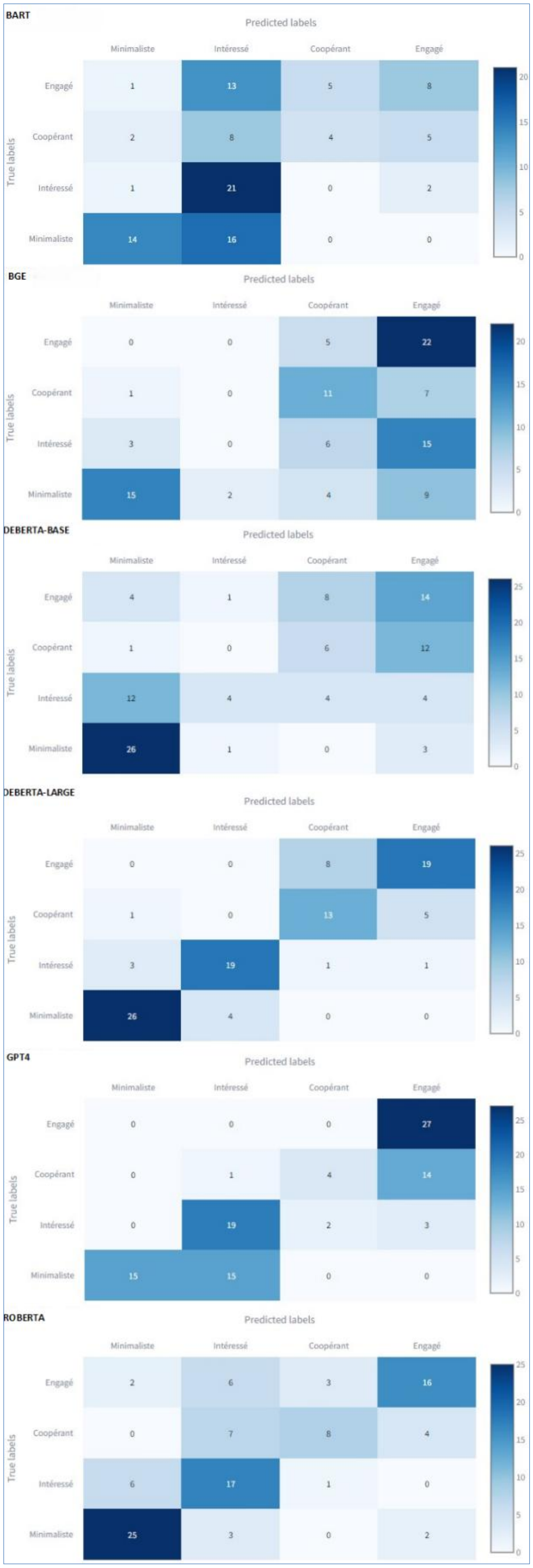

Fig. 6. Confusion matrices for LLM classification results of Task 1.

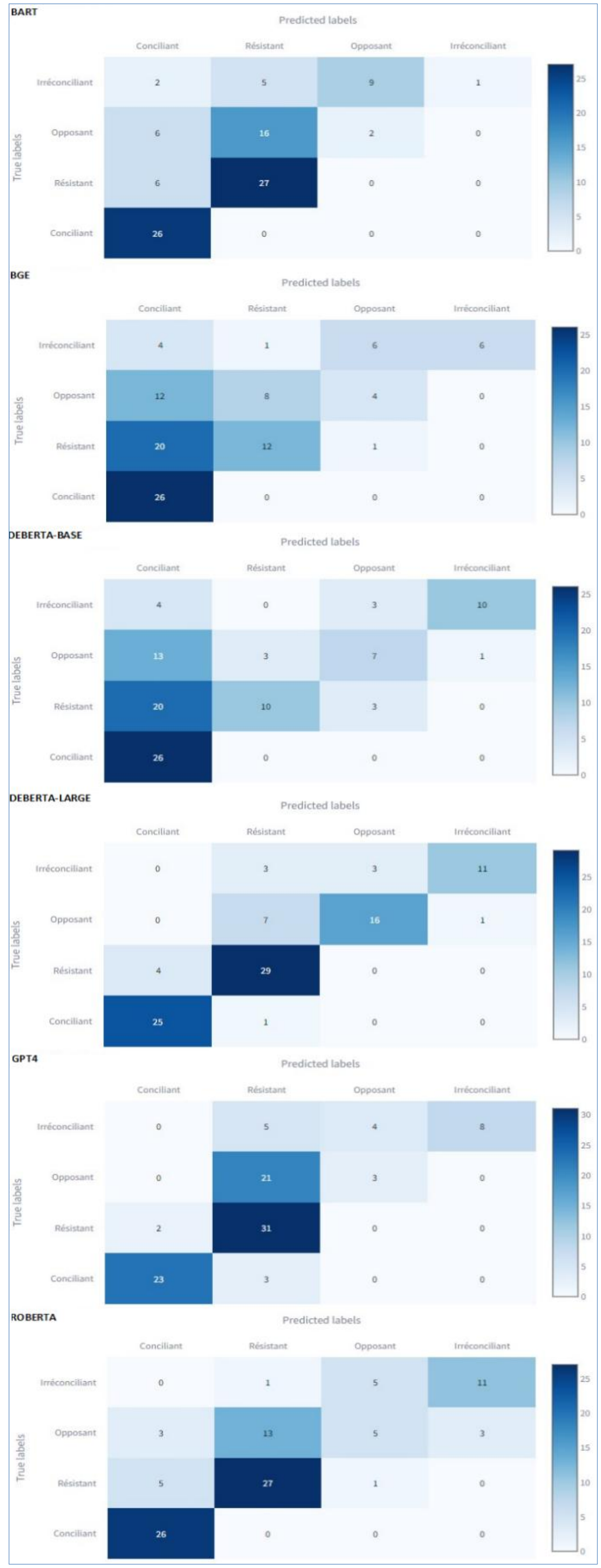

Fig. 7. Confusion matrices for LLM classification results of Task 2.

The results show that the DeBERTa-v3-large-zeroshot model stands out as the best model in terms of F1-score and precision for both synergy classification and antagonism classification tasks, particularly under the explicit prompt. This reflects a good balance between precision and recall. In comparison with modern language models like OpenAI/GPT-4, DeBERTa-v3-large-zeroshot achieves more robust results on these tasks, suggesting that models specifically designed for zero-shot tasks can surpass models like OpenAI/GPT-4 in certain classification contexts without requiring additional fine-tuning. GPT-4, on the other hand, although slightly behind in terms of F1-score, boasts significantly higher precision. This means it generates fewer false positives, demonstrating its ability to minimize false predictions and better handle errors. In addition, RoBERTa-large-zeroshot has a slightly higher F1-score than GPT-4, but lower precision. GPT-4 therefore, remains more accurate, illustrating a difference in the way these models optimize their classification decisions.

In contrast, the Facebook/bart-large-mnli and bge-m3-zeroshot-v2.0 models show significantly lower performance in F1-score and precision, indicating greater difficulty in correctly capturing the right classes and reducing classification errors.

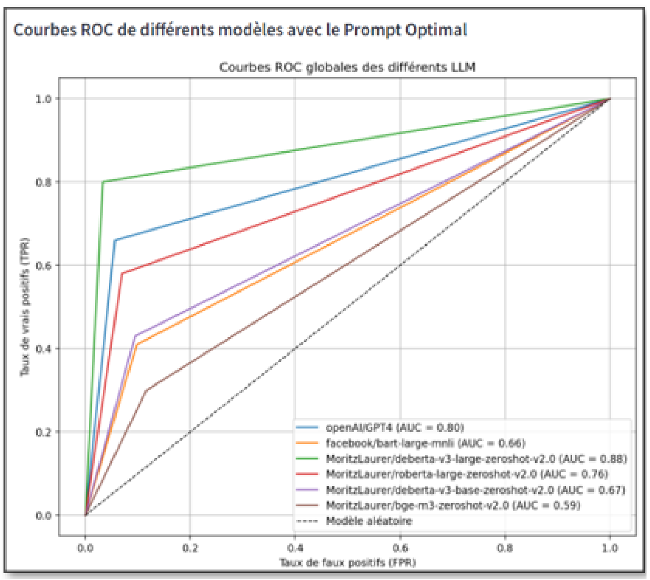

Fig. 8. ROC curves for LLMs.

These discussed findings are confirmed via a review of the Receiver Operating Characteristic Area Under the Curve (ROC AUC) measurements, which gauge the ability of the models to separate positive from negative classifications successfully (see Fig. 8). This analysis reveals that GPT-4 (AUC = 0.80) offers a better discrimination capacity than RoBERTa-large-zeroshot (AUC = 0.76). In fact, although its F1-score is lower, its high precision enables it to achieve a higher AUC, testifying to greater robustness in class distinction. However, DeBERTa-v3-large-zeroshot (AUC = 0.88) outperforms all other models, confirming that it simultaneously maximizes both sensitivity and specificity, making it the top-performing model in classification. In contrast, models like Facebook/bart-large-mnli (AUC = 0.66) and MoritzLaurer/bge-m3-zeroshot-v2.0 (AUC = 0.59) show weaker performance, with behavior closer to a random model in their ability to differentiate classes correctly. This is also reflected in their lower F1-scores and precision, indicating difficulties in accurately classifying positive classes.

VII. Conclusion and Perspectives

This study proposes a new approach for analyzing resistance to change in the context of digital transformation, relying on zero-shot language models (LLMs) and prompting techniques. Unlike traditional methods, which mainly use closed-ended questionnaires and binary classifications, limiting the expression and diversity of employee perspectives, our approach leverages the allies strategy to provide a more granular segmentation of employees' stances towards change. It does so by using open-ended responses for a more detailed and nuanced analysis. The use of specialized zero-shot LLMs in natural language inference, combined with specific prompts, allows for the automatic classification of these open-ended responses without the need for prior training on specific data.

To evaluate the performance of our methodology, we conducted an experiment on an annotated dataset consisting of a hundred employee responses to analyze the models' ability to accurately distinguish the expressed attitudes. Experimental results highlighted the importance of prompt formulation in improving classification accuracy. Specifically, we demonstrated that using explicit prompts optimizes model performance, particularly DeBERTa-v3-large-zeroshot, which emerged as the best-performing model in terms of F1-score and discriminatory power (AUC), surpassing GPT-4. These results also emphasize that, while generalist models like OpenAI/GPT-4 remain competitive, models specifically designed for zero-shot tasks can better meet the requirements of fine classification without requiring additional fine-tuning.

These findings open up several avenues for improvement. One avenue would be to refine the prompting instructions to better capture the nuances of employee responses, thereby optimizing model performance. Additionally, exploring hybrid approaches that combine zero-shot learning with supervised learning could further enhance the classification of stances toward change. Beyond textual analysis, one major research hurdle is blending in various assessment methods, like responses that rely on spoken language. By allowing employees to provide spoken feedback, it is possible to improve the richness of the input data while making the assessment process more inclusive for people with limited writing skills. Finally, applying this methodology to larger and more diverse datasets would allow for evaluating its generalization to different organizational contexts.

Thus, this study demonstrates that zero-shot LLMs, coupled with prompting techniques and the allies strategy, represent a step forward in automating the assessment of resistance to change, providing businesses with a more effective decision-making tool in supporting their digital transformations.

References

[1] B. M. DuBose and A. M. Mayo, “Resistance to change: A concept analysis,” Nurs Forum, vol. 55, no. 4, pp. 631–636, Nov. 2020, doi: 10.1111/nuf.12479.

[2] S. Oreg, "Personality, context, and resistance to organizational change," European Journal of Work and Organizational Psychology, vol. 15, no. 1, pp. 73–101, 2006, doi: 10.1080/13594320500451247.

[3] M. Asem, "Study and synthesis of concepts and practices of resistance to organizational change," Economica, vol. 123, no. 1, pp. 45–55, 2023.

[4] H. Chammaa, R. Ed-Daoudi, and K. Benazzi, "Large Language Models for Academic Internal Auditing," ijacsa, vol. 16, no. 1, 2025, doi: 10.14569/IJACSA.2025.0160166.

[5] V. Cieslak and C. Valor, "Moving beyond conventional resistance and resistors: an integrative review of employee resistance to digital transformation," Cogent Business & Management, vol. 12, no. 1, p. 2442550, Dec. 2025, doi: 10.1080/23311975.2024.2442550.

[6] A. Blyukov, "Resistance to Digital Transformation Projects in FMCG Companies," Scientific Research and Development. Russian Journal of Project Management, vol. 10, pp. 3–12, Aug. 2021, doi: 10.12737/2587-6279-2021-10-2-3-12.

[7] S. Yun, G. Penghe, W. Sitong, and D. R. B. Rosli, "The Impact of Digital Transformation on Business Strategy," vol. 5, no. 6, 2023.

[8] H. G. DEMİREL, "Why do digital transformation initiatives fail? digital mindsets as an invisible strategic component," Yönetim Bilimleri Dergisi, 2024, doi: 10.35408/comuybd.1516665.

[9] B. M. Omowole, A. Q. Olufemi-Phillips, O. C. Ofodile, N. L. Eyo-Udo, and S. E. Ewim, "Barriers and drivers of digital transformation in SMEs: A conceptual analysis," International journal of scholarly research in science and technology, vol. 5, no. 2, pp. 019–036, 2024, doi: 10.56781/ijsrst.2024.5.2.0037.

[10] S. Syamsuddin, S. Marsudi, B. Hasanuddin, A. Umar, and D. Suprayitno, "Adapting to Digital Transformation: Challenges and Strategies for Traditional Businesses," Global, vol. 2, no. 3, pp. 704–711, Mar. 2024, doi: 10.59613/global.v2i3.121.

[11] A. Scholkmann, "Resistance to (Digital) Change: Individual, Systemic and Learning-Related Perspectives," 2021. doi: 10.1007/978-3-030-55878-9_13.

[12] S. P. Robbins and T. Judge, Organizational behavior. Pearson South Africa, 2009.

[13] H. Karaxha, "Methods for Dealing with Resistance to Change," Baltic Journal of Real Estate Economics and Construction Management, vol. 7, no. 1, pp. 290–299, Jan. 2019, doi: 10.2478/bjreecm-2019-0018.

[14] T. Kuzhda, "Diagnosing resistance to change in the change management process," Econ. manag. sustain., vol. 1, no. 1, pp. 49–59, Dec. 2016, doi: 10.14254/jems.2016.1-1.5.

[15] H. Samia, "Digital transformation: A development strategy that faces civil servants resistance to changing working conditions : 'Case of public administrations in the Meknes-Fez region,'" vol. 7, Apr. 2024.

[16] S. Oreg, "Resistance to change: Developing an individual differences measure.," Journal of applied psychology, vol. 88, no. 4, p. 680, 2003.

[17] H. Sherif, "The impact of Resistance to Change to Digital Transformation as a Mediating Variable on the Relationship between Training and Performance Efficiency (Applied to the National Organization for Social Insurance in Egypt)," Jun. 2024.

[18] K. R. Stam, J. M. Stanton, and I. R. Guzman, "Employee resistance to digital information and information technology change in a social service agency: A membership category approach," 2004.

[19] B. A. El Karam, R. Marghoubi, F. Khanboubi, and D. Allaki, "A New Approach for the Analysis of Resistance to Change in the Digital Transformation Context," presented at the International Conference of Innovation and New Trends in Information Technology, Springer, 2022, pp. 153–167.

[20] F. Khanboubi and A. Boulmakoul, "Fuzzy intuitionist approach for resistance to change analysis in a digital transformation process," presented at the 2021 International Conference on Decision Aid Sciences and Application (DASA), IEEE, 2021, pp. 957–961.

[21] J.-C. Fauvet and J.-R. Fourtou, La sociodynamique: concepts et méthodes. Les Editions d'organisation, 1996.

[22] J.-C. Fauvet and J.-R. Fourtou, L'élan sociodynamique. Ed. d'Organisation, 2004.

[23] J. D. Pincus, "Employee engagement as human motivation: Implications for theory, methods, and practice," Integrative Psychological and Behavioral Science, vol. 57, no. 4, pp. 1223–1255, 2023.

[24] Y. E. Rachmad, Network Influence Theory. Mashhad Ferdowsi Ketab Nashriyat, Mahsoos Nashriyat 2023., 2023.

[25] P. Walley, "Stakeholder management: the sociodynamic approach," International Journal of Managing Projects in Business, vol. 6, no. 3, pp. 485–504, 2013, doi: 10.1108/IJMPB-10-2011-0066.

[26] Institut de la Sociodynamique, "Le modèle de la stratégie des alliés." Accessed: Mar. 15, 2025. [Online]. Available: https://www.institutdelasociodynamique.com

[27] P. Bélorgey and N. Van Laethem, "Outil 4. La sociodynamique pour engager les parties prenantes," BàO La Boîte à Outils, vol. 2, pp. 46–54, 2023.

[28] D. Autissier, J.-M. Moutot, K. Johnson, and E. Metais-Wiersch, "Outil 19. La cartographie des alliés," BàO La Boîte à Outils, vol. 2, pp. 58–59, 2022.

[29] A. Vaswani et al., "Attention is all you need," Advances in neural information processing systems, vol. 30, 2017.

[30] S. A. Alsaedi, A. Y. Noaman, A. A. A. Gad-Elrab, F. E. Eassa, and S. Haridi, "Leveraging Large Language Models for Automated Bug Fixing," ijacsa, vol. 15, no. 12, 2024, doi: 10.14569/IJACSA.2024.0151247.

[31] Z. Wang, Y. Pang, and Y. Lin, "Large Language Models Are Zero-Shot Text Classifiers," Dec. 02, 2023, arXiv: arXiv:2312.01044. doi: 10.48550/arXiv.2312.01044.

[32] T. Schopf, D. Braun, and F. Matthes, "Evaluating Unsupervised Text Classification: Zero-shot and Similarity-based Approaches," in Proceedings of the 2022 6th International Conference on Natural Language Processing and Information Retrieval, Dec. 2022, pp. 6–15. doi: 10.1145/3582768.3582795.

[33] S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning, "A large annotated corpus for learning natural language inference," Aug. 21, 2015, arXiv: arXiv:1508.05326. doi: 10.48550/arXiv.1508.05326.

[34] D. O. Romero Mogrovejo, E. Agirre Bengoa, and O. López de Lacalle Lecuona, "Natural Language Inference Models for Few-Shot Text Classification: A Real-World Perspective," 2023.

[35] A. Williams, N. Nangia, and S. R. Bowman, "A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference," Feb. 19, 2018, arXiv: arXiv:1704.05426. doi: 10.48550/arXiv.1704.05426.

[36] Y. Nie, A. Williams, E. Dinan, M. Bansal, J. Weston, and D. Kiela, "Adversarial NLI: A New Benchmark for Natural Language Understanding," May 06, 2020, arXiv: arXiv:1910.14599. doi: 10.48550/arXiv.1910.14599.

[37] J. Thorne, A. Vlachos, C. Christodoulopoulos, and A. Mittal, "FEVER: a large-scale dataset for Fact Extraction and VERification," Dec. 18, 2018, arXiv: arXiv:1803.05355. doi: 10.48550/arXiv.1803.05355.

[38] T. Brown et al., "Language models are few-shot learners," Advances in neural information processing systems, vol. 33, pp. 1877–1901, 2020.

[39] J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," Advances in neural information processing systems, vol. 35, pp. 24824–24837, 2022.

[40] A. Kong et al., "Better Zero-Shot Reasoning with Role-Play Prompting," Mar. 14, 2024, arXiv: arXiv:2308.07702. doi: 10.48550/arXiv.2308.07702.

[41] J. Liu et al., "Generated Knowledge Prompting for Commonsense Reasoning," Sep. 28, 2022, arXiv: arXiv:2110.08387. doi: 10.48550/arXiv.2110.08387.

[42] R. Ghilene, D. Niaouri, M. Linardi, and J. Longhi, "Analysis of Socially Unacceptable Discourse with Zero-shot Learning," Sep. 10, 2024, arXiv: arXiv:2409.13735. doi: 10.48550/arXiv.2409.13735.

[43] T. Schick and H. Schütze, "Exploiting Cloze Questions for Few Shot Text Classification and Natural Language Inference," Jan. 25, 2021, arXiv: arXiv:2001.07676. doi: 10.48550/arXiv.2001.07676.

[44] L. Reynolds and K. McDonell, "Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm," Feb. 15, 2021, arXiv: arXiv:2102.07350. doi: 10.48550/arXiv.2102.07350.

[45] "Hugging Face Models." Accessed: Dec. 30, 2024. [Online]. Available: https://huggingface.co/models

[46] M. Lewis et al., "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension," Oct. 29, 2019, arXiv: arXiv:1910.13461. doi: 10.48550/arXiv.1910.13461.

[47] P. He, X. Liu, J. Gao, and W. Chen, "DeBERTa: Decoding-enhanced BERT with Disentangled Attention," Oct. 06, 2021, arXiv: arXiv:2006.03654. doi: 10.48550/arXiv.2006.03654.

[48] Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," Jul. 26, 2019, arXiv: arXiv:1907.11692. doi: 10.48550/arXiv.1907.11692.

[49] J. Achiam et al., "Gpt-4 technical report," arXiv preprint arXiv:2303.08774, 2023.