

# Developing Machine Learning Models for the Analysis of Citizens' Contributions from E-Participation\*

Marten Borchers<sup>1,\*</sup>, Delzar Habash<sup>2</sup> and Eva Bittner<sup>1</sup>

<sup>1</sup>University of Hamburg, Mittelweg 177, 20148 Hamburg, Germany

<sup>2</sup>Deutsches Zentrum für Luft- und Raumfahrt e.V.

## Abstract

Citizen participation is increasingly relevant in acquiring local knowledge for urban projects. However, the increasing number of participants has created a problem of data overload, as manual analysis is hardly feasible due to its expense, time intensity, and slowness. AI and ML can help reduce or solve this issue. However, due to the lack of proper ML-based approaches, we investigate how textual contributions from citizens can be analyzed using machine learning techniques. To achieve this, we followed the knowledge discovery in the database framework, collected data, and trained several machine learning models, which we analyzed and compared. Our findings demonstrate that urban development contributions often cover multiple topics, making classification challenging, which also corresponds to the length of citizens' contributions. Transformer models, however, show remarkable precision when compared to SVM models. With our findings, we contribute to the analysis of citizen contributions to support democratic processes and scalable citizen participation.

## Keywords

Citizen Participation, Machine Learning, Urban Planning, Data Analysis, Knowledge Discovery in Databases

## 1. Introduction

Today, most of the global human population lives in urbanized areas, and this trend is expected to continue until 2050 [1]. However, this growth has led to difficulties for urban developers and planners, who aim to plan projects in the most efficient way possible, thus saving resources, improving the quality of life [2], and increasing sustainability [3]. Planning such projects comes with its own set of problems [4]. Ideally, the main stakeholders should be the local citizens since they live in the area concerned. However, citizens' wishes may conflict with the priorities and objectives of other stakeholders like local businesses and industries, development or construction companies, environmental and conservation agencies or groups, local government policies, and more [5]. As a reaction, governments involve citizens in urban planning to increase transparency, trust, acceptance, and the real-world benefit of projects [6].

However, electronic participation (e-participation) with websites and comprehensive participation platforms, such as Decide Madrid [7] and DIPAS [8], among others, lacks automatic analysis systems that can handle thousands of citizen contributions to support urban experts in subsequent processes. For example, in 2019, a Scottish portal generated around 16,000 responses [9]. The absence of automatic analysis approaches and the fact that the gathered citizen contributions are mainly in free-text [10] increase costs and errors during the analysis process, e.g., inconsistencies. Therefore, new solutions are required to automatically analyze citizens' contributions [11].

The scientific literature describes demands and requirements, e.g., summarized and focused on sentiment analysis, keyword search, and clustering [12], and presents initial solutions for machine learning models (ML). Natural language processing (NLP) is crucial in handling textual data. Despite the significant strides in the field of NLP, the processing of participation data with NLP, particularly classification, remains understudied, which can be attributed to a lack of publicly available data.

*Proceedings EGOV-CeDEM-ePart conference, August 31 - September 4, 2025, University for Continuing Education, Krems, Austria*

\*Corresponding author.

✉ marten.borchers@uni-hamburg.de (M. Borchers); delzar.habash@dlr.de (D. Habash); eva.bittner@uni-hamburg.de (E. Bittner)

ORCID 0000-0003-4356-8950 (M. Borchers); 0000-0002-7628-6012 (E. Bittner)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Therefore, we build upon the existing knowledge of artificial intelligence (AI), ML, and NLP in citizen participation to examine the following research question (**RQ**).

**RQ:** How should textual citizens' contributions from e-participation be classified with machine learning techniques?

To answer the **RQ**, we pursue a practical and data-oriented approach as described by the Knowledge Discovery in Databases Framework (KDD), which defines a structured process for gaining new knowledge from data and ML models [13]. Therefore, we collected a dataset from e-participation and manually labeled it. We then trained several ML models, which were compared to each other and further analyzed to increase their accuracy, as described and elaborated on in the following sections.

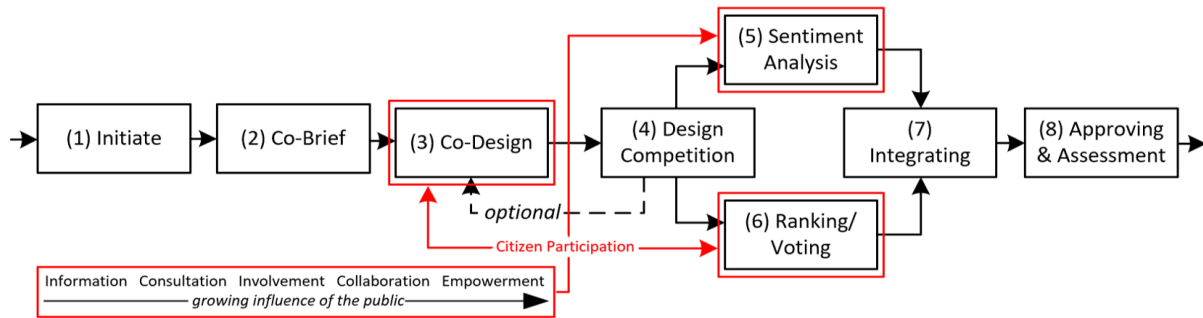
## 2. Theoretical Foundation

Citizen participation can be defined as an involvement characterized by interactive and iterative deliberation processes among citizens and government officials [14]. Another definition, provided by Boudjelida et al., is "a form of democracy in which citizens are part of the decision-making process about the development of their society" [15]. Both show that solutions or views imposed on the public top-down by leaders and experts are no longer considered optimal and that governing institutions need to build networks to link them with the public to keep up with modern complexities to which many agree [16].

### 2.1. Minimum Viable Process for Citizen Participation

Stelzle et al. [17] proposed a process model for incorporating citizens and other stakeholders in urban projects, which has also been tested and applied in practice, called the minimal viable process (MVP), summarized in Figure 1.

The MVP begins with (1) initiating the urban project and assembling a project team. In step (2) (co-brief), the project lead's objectives are documented and prioritized with stakeholders, including designers and planning authorities, focusing on project scope and citizen participation. This sets the stage for (3) co-design, where citizen participation occurs through various methods like information booths, workshops, (online) surveys, contests, social media, and participation platforms and therefore contributes huge amounts of textual data [16].



**Figure 1:** Minimum Viable Process (MVP) for Citizen Participation.

At the end of the co-design process, results are evaluated to determine the project's strategic alignment. In step (4), project visions developed by architects and urban planners are published to gather further feedback from citizens through voting and comments on digital platforms. Steps (5) and (6) involve analyzing comments and voting results, which are then integrated (7) into a comprehensive urban project design. Finally, this design must be approved (8) by the project lead and planning authorities (9) to ensure compliance with regulations and facilitate permit acquisition [17].

## 2.2. Challenges in the Analysis of Citizens' Contributions

Citizens' inputs generated during participation must be analyzed and evaluated by experts (cf. Figure 1) [12]. Experts face challenges in this context, as the number of participants in e-participation and on participation platforms has increased significantly [9]. This creates an overload of information for subsequent processes, including analysis, sense-making, and the development of shared ideas and visions [18].

A larger number of participants is beneficial for e-participation because it generates broader, more diverse, and more comprehensive knowledge [19]. However, massive participation events, such as the U.S. Federal Communications Commission's net neutrality implementation, which gathered over 450,000 comments [20], showcase how the volume of contributions can increase. Experts often lack the resources and skills to exploit this knowledge pool [21]. The extensive data volumes pose significant challenges for processing, and the resulting information overload becomes a real issue. Real-time analysis could significantly reduce the time between analysis and decision-making [22]. Another challenge in analyzing citizens' contributions is the individual and qualitative processing [12]. Evaluators can build on reusable and standardized methods and categories, saving time and financial resources [16].

## 2.3. AI and ML for analyzing E-Participation Data

AI encompasses various technical approaches that facilitate the automatic execution of tasks that mimic human capabilities [23]. Within this domain, ML refers to algorithms that enable models to learn from data to make predictions, classifications, or clustering [24]. NLP is a subfield of AI and ML that focuses specifically on analyzing and understanding human language, making it a valuable tool for interpreting citizen contributions in free-text formats. Despite the increasing interest in leveraging ML and NLP for e-participation initiatives, practical applications remain limited [10].

Recent studies highlight the potential of ML and NLP in addressing the challenge of information overload on e-participation platforms. Arana-Catania et al. [21] proposed utilizing these technologies to analyze citizens' interests, fostering increased participation and enhancing user experience on urban platforms. Similarly, Lieven et al. [8] emphasized the efficiency of using ML to evaluate complex and unstructured citizen inputs by generating automatic summaries. Moreover, researchers have examined clustering models to analyze sentiments from social media contributions, aiding moderators in real time [25]. However, the ambiguity inherent in short or keyword-based contributions poses challenges for accurate interpretation. Romberg et al. [22] provide an analogous approach, illustrating the training of a classification model with citizen inputs from a mobility project to establish a preliminary analysis framework. These studies collectively underscore the promise of ML and NLP in enhancing the analysis of citizen contributions.

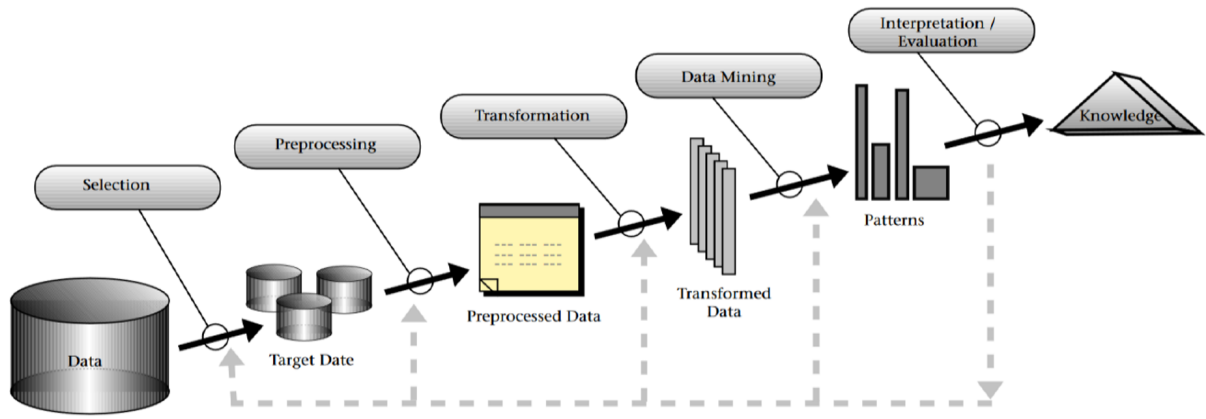
# 3. Research Methodology

As stated by the RQ, ML models have not yet produced sufficient results for analyzing citizens' contributions from e-participation. Given the importance of data, e.g., citizen contributions, we follow the Knowledge Discovery in Databases Framework (KDD) [13].

## 3.1. Knowledge Discovery in Databases Framework

The framework delineates a systematic methodology for iteratively deriving insights or knowledge from data [26]. This framework operates on the premise that IT systems can document data related to usage, user behaviors, and outcomes, which, when analyzed, facilitate the making of inferential statements and a deeper understanding. Furthermore, a robust dataset enables the training of new ML models. Knowledge generated through the framework is expected to be generalizable, non-trivial, novel, useful, and comprehensible [26].

The KDD is distinguished by its structured process, as illustrated in Figure 3.1, and is critically relevant in today's data-driven world [27]. Unlike more pragmatic and flexible methodologies focused



**Figure 2:** Knowledge Discovery in Databases Framework [26].

on resolving specific business challenges, the KDD emphasizes a rigorous protocol that includes data cleansing and standardization, thereby facilitating the replication and verification of results by other researchers. In the initial phase, the goal and domain knowledge are defined. Subsequently, the KDD procedure is traversed in five phases (cf. Figure3.1). In the selection phase, the data suitable for the goal are chosen. In the preprocessing phase, data quality is evaluated and cleaned. The data is converted into a format that ML algorithms can process in the transformation phase. In the data mining or ML phase, algorithms are selected to search for patterns or insights in the data [26]. The KDD process focuses on the data, which is why ML or NLP can be employed. In the interpretation and evaluation phase, the results of the ML algorithm are finally analyzed and visualized [13].

### 3.2. Application of the Framework

In applying the framework, we elicited a dataset of citizen contributions to urban planning in Germany. We archived that by combining available data from two past participations. One was implemented in eastern Germany, and the other in middle Germany. Both focused on the reconstruction of buildings: the first was a university campus in a larger city, and the second was a school in a smaller city. Together, the dataset comprises nearly 18,000 contributions from citizens.

**Table 1**  
Overview of the Dataset.

No.	Class	Entries	Proportion
1.	Mobility	6,059	35.85 %
2.	Environment	2,409	14.14 %
3.	Living	1,813	10.65 %
4.	Social Network	1,313	7.78 %
5.	Education	1,222	7.718 %
6.	Economy	1,145	6.72 %
7.	Culture	536	3.15 %
8.	Sport	370	2.17 %
9.	Public Service	206	1.21 %
10.	Undefined	1,946	11.43 %
Sum		17,031	100.00 %

In the selection step, we conducted a detailed examination of the data set to understand its nature and properties. This preliminary analysis is critical, laying the groundwork for all subsequent stages of the framework process. In the preprocessing step, we cleaned the data, addressed irregularities, and

handled missing values by manually correcting or removing them to create a uniform dataset of 17,031 entries as shown in Table 1.

The transformation step was extensive. To train a classification model, we had to label all entries. This was achieved with the help of two hired and trained students who labeled all data manually according to the ten categories as shown in Table 1. Both students labeled the data according to a description of the categories in the literature [12, 22]. The last category was used to count unclear and not understandable contributions. The prepared data was formatted appropriately for text categorization, involving organizing the texts and their corresponding labels into a structured dataset.

In the data mining step, we conduct various experiments to study the data characteristics and identify potential challenges that impede effective text categorization. These experiments also helped in understanding the baseline performance of NLP techniques on the data [24]. In the evaluation and interpretation step, we focused on the performance of the developed ML models and interpreted their effectiveness in classifying contributions within citizen participation. ML Prototypes were tested, and an ensemble approach was favored due to the unbalanced and multi-topic nature of our dataset [28].

## 4. Evaluation and Findings

The dataset under investigation was partitioned into a training set (80%) and a testing set (20%) and was subjected to classification using five different classifiers, as shown in Table 2. The dataset underwent several sequential transformations to investigate the effects of the preprocessing techniques on the performance of successive transformations. Each subsequent transformation builds upon the output of the previous one.

The initial condition involved no preprocessing. Therefore, it contained only the pure text, vectorized using the wiki corpus dataset [29]. The next step involved normalizing capitalization so that the holder text is in lowercase to increase semantic similarity, and the last step was to stem the tokens, ensuring that words are converted into their base form. The performance of the classifiers was evaluated using the F1 score, a harmonic mean of precision and recall calculated via the confusion matrix (true-positive, false-positive, true-negative, and false-negative), which is particularly useful in situations where a balance between these metrics is desired [30]. This measure is crucial when dealing with datasets that may have an imbalance in class distribution or when the costs of false positives and false negatives are significant.

**Table 2**  
Comparison of Trained ML Models.

No.	Model with F1-Score(s)	Pure Text	Without Stop Words	With Capitals	Stemmed
1.	Naive Bayes	0.44	0.48	0.48	0.49
2.	Logistic Regression	0.67	0.66	0.66	0.67
3.	Support Vector Machine	0.69	0.68	0.68	0.69
4.	Decision Tree Classifier	0.50	0.56	0.56	0.58
5.	Gradient Boosting	0.62	0.61	0.61	0.63

In the training and analysis conducted, support vector machine (SVM) and logistic regression exhibited superior performance. This robust performance was observed across various stages of text preprocessing. Naive Bayes showed notable improvement when stop words were removed, consistent with their probabilistic foundations, which can be sensitive to irrelevant features. Decision trees, which are generally weaker classifiers for text data but robust against outliers, showed poor performance across all text variants, whereas ensemble methods, such as gradient boosting, demonstrated enhanced outcomes.

Further analysis was conducted using the same classifiers as before. Similar preprocessing steps were applied except for removing special characters, which had already been conducted during data cleaning. SVM again performed the best among the classifiers, though the F1 scores indicated that the classifiers struggled more with citizens' contributions than with the initial dataset. Attempts to optimize the

SVM's hyperparameters did not yield significant improvements, with the F1 score plateauing around 0.7. A fascinating insight was gained by excluding entries labeled as "undefined," which typically represent contributions that are too vague or ambiguous to be classified meaningfully. This exclusion led to a noticeable improvement in the SVM's performance, from an F1 score of 0.69 to 0.75. This improvement underscores the importance of dataset quality and the impact of ambiguous data on classifier performance.

The trained SVM models concentrated on the length of citizens' contributions to identify possible patterns that ML models could abstract. The existing literature suggests that longer texts generally enhance classifier performance. However, this assumption requires validation within the domain of citizen participation [28], where texts are typically quite short, nuances differ significantly, and empirical data are limited. This validation involved conducting three different approaches to assess how text length affects classification in this context. The data processing was conducted as before, and training and test data were the same size as before. Notably, while previous studies measured text length by character count or word count, this analysis used word count [31].

**Table 3**

Overview of F1-Score for Citizen's Maximum Length of Contribution for the trained Support Vector Machine Model.

Max. Word Count	2	3	4	5	6	7	8	9	10
n-Entries	937	1,931	3,075	4,144	5,175	6,155	6,999	7,805	8,490
F1-Score	0.708	0.750	0.771	0.761	0.760	0.760	0.750	0.748	0.748

Table 3 shows that the length of citizens' contributions influences the performance of the trained SVM models. In columns 4 to 7, an increase in performance was observed with longer text lengths, surpassing baseline levels and corroborating previous research [12]. However, benefits from increasing text length plateaued after reaching eight words, with slight declines in performance thereafter. While these findings are not definitive due to the variable dataset sizes, the trend becomes more apparent when considering the results together, further suggesting that longer texts are leading to diminished performance.

**Table 4**

Overview of F1-Score for Citizen's Minimum Length of Contribution for the trained Support Vector Machine Model.

Min. Word Count	2	3	4	5	6	7	8	9
n-Entries	14,859	13,151	12,007	10,938	9,907	9,827	8,083	7,277
F1-Score	0.733	0.730	0.724	0.714	0.708	0.703	0.697	0.696

Our findings indicated an unexpected trend: as text length increased, the F1 score generally decreased, suggesting that longer texts might not necessarily yield better models. This was contrary to prior assumptions [22] and raises questions about the optimal text length for training classifiers in citizen participation. However, longer texts can also lead to poorer performance when the narrative deviates from the central theme, complicates label assignment, or tends to be ambiguous. This pattern led to the conception of the third training of SVM models, where a sliding text length window was applied to the dataset to fine-tune the optimal word count range. The results, presented in Table 4, begin with an evaluation of model performance in identifying a specific range where text length maximizes classification accuracy.

Our analysis indicated that datasets with shorter texts, as displayed in the first column of Table 5, exhibited the highest performance. Performance consistently declined as the text length window expanded to include longer texts. This pattern was not attributable to dataset size, as datasets with shorter texts performed significantly better than those with longer texts of comparable size, identifying text length as the critical variable affecting model performance. These three approaches collectively suggest



**Table 5**

Overview of F1-Scores for different Word Windows in Number of Words for the trained Support Vector Machine Model.

Word Windows	1-5	2-6	3-7	4-8	5-9	6-10	7-11	8-12
n-Entries	3,075	4,137	4,952	5,218	5,068	4,730	4,346	3,905
F1-Score	0.772	0.764	0.759	0.747	0.733	0.701	0.695	0.672

that the established research findings might not directly apply to the domain of citizen participation data, e.g., citizens' contributions. Although evidence was found that shorter contributions are easier to categorize, variations could arise with different datasets and methodologies. This insight is valuable for facilitators, who may encourage participants to make brief contributions for more efficient processing.

**Table 6**

Overview of F1-Scores for trained Models with the same Number of Entries per Class for the trained Support Vector Machine Model.

Entries per Class	400	500	600	700	800	900	1,000	1,100	1,145
n-Entries	2,400	3,000	3,600	4,200	4,800	5,400	6,000	6,600	6,870
F1-Score	0.704	0.723	0.741	0.758	0.772	0.784	0.792	0.802	0.806
Improvement	-	1.9 %	1.8 %	1.7 %	1.4 %	1.2 %	0.8 %	1.0 %	0.4 %

To examine the impact of dataset size and the number of entries per class, we trained several additional ML models. The F1 score of SVM models that contained at least 1,145 entries is shown in Table 6. The classes included are Mobility, Environment, Living, Social Network, Education, and Economy (cf. Table 1). The performance increased significantly when the dataset size increased over the trained models up to 1,145 entries per class. Although results kept improving with each increment, diminishing returns were noticeable, as denoted by the last row of Table 6. This suggests a performance plateau at a certain dataset size, but the dimensions of the plateau could not be captured with the limited amount of the citizens' contributions.

**Table 7**

Overview of F1-Scores for the upscaled Transformer Model.

Dataset	Entries	Classes	Duplicates	Duplicate Proportion	F1-Score
1.	1,145	6	False	0 %	0.79
2.	1,200	6	True	3 %	0.86
3.	1,500	6	True	4 %	0.91
4.	2,000	6	True	5 %	0.92

Next, we focused on transformer technology. We trained a transformer model (DistilBERT) using default parameters from the Simple Transformers library (Python), with a batch size of 8 and over 8 epochs, achieving an F1 score of 0.759% on an evaluation dataset of 600 instances. Further trained models, including an up-sampling of the training data with duplicates, were allowed, and datasets were assessed at 1,200, 1,500, and 2,000 instances per class (cf. Table 1). Despite maintaining less than 5% duplicates in test sets to minimize bias, performance increased impressively, highlighting that transformer models benefit from larger training datasets. Incremental epochs 12 and 16 showed no significant improvement, reinforcing the model's responsiveness to dataset size over extended training durations. However, an accurate classification with an F1 score of 0.92, as shown in Table 6, is promising, even if it was only achieved with the six classes of Mobility, Environment, Living, Social Network, Education, and Economy (cf. Table 1) due to a lack of further data.

## 5. Discussion

In this paper, we examined the **RQ** "How should textual citizens' contributions from e-participation be classified with machine learning techniques?" by applying the KDD Framework, which provided a structural approach for analyzing data and deriving knowledge. To answer the **RQ**, we collected and created a labeled dataset of 17,031 preprocessed entries representing real citizens' contributions from past participation projects, which we used to train and analyze multiple ML models.

For practitioners and scientists, the dataset created can serve as a baseline for future development and currently represents the largest dataset so far, which we hope will soon be extended. This desire is also reflected in our findings. The trained ML models demonstrate that training is possible with an acceptable F1 score, e.g., validity, especially for the classes in which our dataset contains many entries. However, as urban experts and decision-makers should develop a proper vision and decide if these sometimes very costly (construction) projects are implemented, a high validity of the analysis of data is required, leading to the following insights.

**Insight 1:** Only validated ML models with high validity should be used to analyze citizens' contributions from e-participation, which requires more data.

**Insight 2:** Preprocessing is mandatory, especially regarding the meaning of contributions and their decomposition into sense-making pieces.

**Insight 3:** ML model classifiers must represent the desired classes by experts in the training data equally to increase the accuracy of the prediction.

**Insight 4:** The length of the text contributions is decisive for the classification and must be preprocessed as required and divided into semantic paragraphs.

Considering insight 1, we cannot say which validity of an F1 score is required. However, the trained transformer models showed the highest scores and a value above 0.9 may be enough to support experts, while a fully automatic analysis might require an even higher validity. Insight 2 is derived from the findings of Table 2, Table 4, Table 5, and Table 6. In creating the dataset, all contributions were divided into sense-making pieces, allowing for abstraction and highlighting existing findings [32]. This is represented in the results and should be considered. Alternatively, it is possible to predict multiple classes at once. However, the abstraction in terms of semantics usually becomes more advanced, and the number of classes is fixed, which reduces the flexibility of the current approach. Insight 3 highlights the demand for an equal proportion of class entries, as shown in Table 7. Furthermore, our current results indicate that shorter texts generally yield better classification results, although this finding may be context-specific. Nevertheless, this, along with the increasing F1 score in Table 7, highlights the need for further data that we specify with Insight 5.

**Insight 5:** ML model classifiers should at least contain 1,145 (preferably more) entries per class for the training to lay a profound base for the classification.

**Insight 6:** Transformer models are more capable of classification than SVM in terms of limited datasets but are black box models and not transparent.

**Insight 7:** Practitioners should deviate between transparency and efficiency of the trained model (in SVM or decision trees) and higher accuracy (e.g., black box models like transformer models).

The training of the SVM and transformer models demonstrated that while SVM models are efficient and require shorter training times, transformer models, though slow and resource-intensive, provided superior classification. This highlights a trade-off between efficiency and performance, influenced by the specific needs and resources of participation facilitators. For practitioners involved in e-participation initiatives, the implementation and development of ML models provide a significant opportunity to enhance the analysis of citizen contributions. These models can automate the processing of large volumes of data, enabling a more efficient and systematic approach to understanding citizens' requirements and sentiments [10]. By employing ML techniques, practitioners can offer insights that drive urban



planning and policy-making processes. However, there is a pressing need to tackle all gained insight to effectively design, implement, and interpret ML models [21]. This will lay the technical foundation for the usage of ML and empower urban experts and practitioners to harness the full potential of ML in analyzing the qualitative complexity of citizen contributions, as described in section 2.3.

For governments and urban planners, understanding how ML can be developed and utilized in the context of citizen participation is essential [22]. The integration of ML can facilitate a more democratic and participatory governance framework where real-time data analysis aids decision-making [33]. Our insight and the developed ML models contribute to that and establish an initial baseline for how ML is deployed in analyzing citizen contributions. However, a critical consideration in deploying ML models is acknowledging existing constraints, particularly concerning the availability of training data and a standardized classification framework (cf. Insight 4). With sufficient quality data and established classes, the performance of ML algorithms is likely to be maintained. To navigate these challenges, governments and practitioners must prioritize collecting and curating diverse datasets that accurately reflect the spectrum of citizens' contributions [22]. Additionally, developing standards for data classification will help streamline analysis procedures and enhance the effectiveness of ML applications in this domain [34].

The findings from this research contribute to the scientific literature by confirming existing patterns and revealing new insights into the nature and complexity of citizen contributions across urban planning contexts. By systematically applying the KDD framework to analyze e-participation data, the study yields new insights into the domain of citizen participation and e-participation, enriching the understanding of how ML techniques can effectively categorize qualitative information [28]. This exploration extends the current knowledge base and emphasizes the importance of integrating ML technology into participatory processes [24].

Moving forward, insights 4 to 6 should be interpreted as a research agenda and call for further action from other researchers to develop ML standards, train extended and capable ML models, and also publish their data and source code. This should include defining best practices for data handling, processing techniques, and model evaluation, as well as the selection of models. Building trust in them is also challenging. For instance, transformer models, such as neural networks, are not transparent, while decision trees can be fully visualized and explained [12].

One of the principal limitations of this study is the lack of ample training data, which constrains the performance of the developed ML models. Without a sufficiently large and representative dataset, the models may struggle to generalize findings effectively, potentially leading to biases in classification outcomes. Moreover, the study indicates the need to explore additional ML approaches beyond the transformer models used. Techniques from generative AI may also hold promise for improving the capabilities of citizen contributions analysis, warranting further research into their applicability and effectiveness [35]. The trust of experts and practitioners in these ML models remains a critical issue that requires further examination. Stakeholders must be assured of the reliability, accuracy, and utility of ML systems when applied to citizen data. Therefore, ongoing dialogues and workshops should be conducted to disseminate findings and cultivate confidence in these technologies. Finally, to ensure user engagement and satisfaction, the design of interfaces built around these ML models should be rigorously tested and evaluated [34]. Feedback from urban experts, moderators, and decision-makers will be essential in constructing intuitive and effective analytical tools that foster effective communication between citizens and policymakers [10].

Looking ahead, we aim to enhance and retrain our current datasets to improve the predictive capabilities of our machine learning (ML) models. Ongoing efforts to collect fresh data and refine existing datasets are essential to ensure that our models remain relevant and effective in responding to the ever-changing dynamics of urban environments. Future research should also prioritize the exploration of additional ML models and the implementation of advanced visualization techniques. These enhancements can increase the interpretability of model outputs, building trust and acceptance among stakeholders while effectively conveying insights gathered from citizen contributions. Ultimately, investigating the efficiency potential that ML offers to governments and companies in e-participation can provide significant benefits. By automating data analysis and accelerating decision-making processes,

ML models can promote a more collaborative and practical approach to urban planning, ultimately benefiting both the public and governmental entities.

## 6. Conclusion

In this paper, we examined the **RQ** "How should textual contributions from citizens be analyzed using machine learning techniques to classify e-participation data?" The efforts undertaken in this paper aimed to find the appropriate methods and techniques for combining the technical aspects of modern ML paradigms with established research on citizen participation.

After thoroughly examining the relevant research on e-participation, its current implementations, advantages, and weaknesses, it became clear that scholarly deficiencies exist. Participation facilitators, decision-makers, and official bodies often lacked the necessary tools to process, aggregate, and draw unbiased conclusions from participation data. The lack of will and resources to conduct meaningful participation, among other factors, leads to apathy and distrust among the public. The theoretical field of classical participatory democracy is well-studied, and the barriers to participation are known. However, the scientific knowledge base to overcome technical barriers that hinder the advancement of participatory democracy is yet far from sufficient. Efforts to facilitate mass participation by creating adequate IT artifacts are lacking.

The suggested models can extract the topic, e.g., the class of any contribution from a set of defined classes, to an impressive degree of accuracy. The presented SVM ensemble can also extract all present topics in a contribution, thus performing multi-class classification. This gives participants the freedom to express as many wishes as desired in their contributions, and it allows decision-makers to understand where the public interest lies, enabling them to create better plans that are satisfactory for most citizens. However, in addition to the trained ML models and findings about the characteristics of citizens' contributions, further research is needed to enhance accuracy and expand existing and publicly available datasets. Overall, while the results are promising, they underscore the need for further research to explore the nuances in citizens' contributions to developing accurate ML classifiers.

## Acknowledgments

We want to express our sincere gratitude to the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) for its funding of the RESCUE-MATE project. The project examines innovative approaches to support citizen participation and civic engagement in urban planning and emergency management. The funding of RESCUE-MATE, with the grant numbers 13N16835 to 13N16845, enabled us to carry out this research and this paper.

## Declaration on Generative AI

During the preparation of this work, the author(s) used the GPT-4 mini and GPT 4.1 models from Open AI to translate texts and to support the writing and development process in Python. Additionally, we utilize Grammarly for spelling checks. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

## References

- [1] United Nations, World cities report 2022, 2022. URL: <https://unhabitat.org/wcr/>, last accessed 2024-10-28.
- [2] B. R. Barricelli, G. Fischer, D. Fogli, A. Morch, A. Piccinno, S. Valtolina, Copda 2022—cultures of participation in the digital age: Ai for humans or humans for ai?, in: Proceedings of the 2022

International Conference on Advanced Visual Interfaces, Association for Computing Machinery, New York, NY, USA, 2022, pp. 1–3.

- [3] N. Bracarense, R. Bawack, S. Wamba, K. Carillo, Artificial intelligence and sustainability: A bibliometric analysis and future research directions, *Pacific Asia Journal of the Association for Information Systems* 14 (2022).
- [4] K. Axelsson, M. Granath, Stakeholders' stake and relation to smartness in smart city development: Insights from a swedish city planning project, *Government Information Quarterly* 35 (2018) 693–702.
- [5] S. Philpot, K. Hipel, P. Johnson, Identifying potential conflict in land-use planning using a values-centered e-participation tool: A canadian case study in aggregate mining, in: *Hawaii International Conference on System Sciences 2019 (HICSS-52)*, 2019.
- [6] P. Repette, J. Sabatini-Marques, T. Yigitcanlar, D. Sell, E. Costa, The evolution of city-as-a-platform: Smart urban development governance with collective knowledge-based platform urbanism, *Land* 10 (2021) 33.
- [7] I. Cantador, A. Bellogín, M. Cortés-Cediel, O. Gil, Personalized recommendations in e-participation: offline experiments for the “decide madrid” platform, in: *Proceedings of the International Workshop on Recommender Systems for Citizens*, Association for Computing Machinery, New York, NY, USA, 2017, pp. 1–6.
- [8] C. Lieven, B. Lüders, D. Kulus, R. Thoneick, Enabling digital co-creation in urban planning and development, in: A. Zimmermann, R. Howlett, L. Jain (Eds.), *Human Centred Intelligent Systems*, Springer, Singapore, 2021, pp. 415–430.
- [9] S. Royo, V. Pina, J. Garcia-Rayado, Decide madrid: A critical analysis of an award-winning e-participation initiative, *Sustainability* 12 (2020) 1674.
- [10] J. Romberg, T. Escher, Making sense of citizens' input through artificial intelligence: A review of methods for computational text analysis to support the evaluation of contributions in public participation, *Digital Government: Research and Practice* 5 (2024) 3:1–3:30.
- [11] D. Bawden, L. Robinson, Information overload: An overview, 2020. Presented at Oxford, June 1, 2020.
- [12] M. Borchers, T.-B. Cao, N. Tavanapour, E. Bittner, Designing ai-based systems to support the analysis of citizens' inputs from e-participation, in: *ECIS 2024 Proceedings*, 2024.
- [13] O. Maimon, L. Rokach, Introduction to knowledge discovery in databases, in: O. Maimon, L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook*, Springer US, Boston, MA, 2005, pp. 1–17.
- [14] H. Vasudavan, S. Balakrishnan, The taxonomy of smart city core factors, in: *Proceedings of the 2019 7th International Conference on Information Technology: IoT and Smart City*, Association for Computing Machinery, New York, NY, USA, 2019, pp. 509–513.
- [15] A. Boudjelida, S. Mellouli, J. Lee, Electronic citizens participation: Systematic review, in: *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance*, Association for Computing Machinery, New York, NY, USA, 2016, pp. 31–39.
- [16] M. Hofmann, S. Münster, J. Noennig, A theoretical framework for the evaluation of massive digital participation systems in urban planning, *J geovis spat anal* 4 (2020) 3.
- [17] B. Stelzle, A. Jannack, J. R. Noennig, Co-design and co-decision: Decision making on collaborative design platforms 112 (2017) 2435–2444. doi:<https://doi.org/10.1016/j.procs.2017.08.095>.
- [18] T. Glatz, M. Lippold, Is more information always better? associations among parents' online information searching, information overload, and self-efficacy, *International Journal of Behavioral Development* 47 (2023) 444–453.
- [19] A. Panori, C. Kakderi, N. Komninos, K. Fellnhöfer, A. Reid, L. Mora, Smart systems of innovation for smart places: Challenges in deploying digital platforms for co-creation and data-intelligence, *Land Use Policy* 111 (2021) 104631.
- [20] P. Toor, K. Olmstead, Skye, Public comments to the federal communications commission about net neutrality contain many inaccuracies and duplicates, 2017. Last accessed 2024-06-15.

- [21] M. Arana-Catania, F.-A. Lier, R. Procter, N. Tkachenko, Y. He, A. Zubiaga, M. Liakata, Citizen participation and machine learning for a better democracy, *Digital Government: Research and Practice* 2 (2021) 27:1–27:22.
- [22] J. Romberg, L. Mark, T. Escher, A corpus of german citizen contributions in mobility planning: Supporting evaluation through multidimensional classification, in: *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, European Language Resources Association, Marseille, France, 2022, pp. 2874–2883.
- [23] D. Susar, V. Aquaro, Artificial intelligence: Opportunities and challenges for the public sector, in: *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance*, Association for Computing Machinery, New York, NY, USA, 2019, pp. 418–426.
- [24] M. Cai, Natural language processing for urban research: A systematic review, *Heliyon* 7 (2021) e06322.
- [25] V. Pitogo, C. Ramos, Social media enabled e-participation: a lexicon-based sentiment analysis using unsupervised machine learning, in: *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 518–528.
- [26] U. Fayyad, Knowledge discovery in databases: An overview, in: N. Lavrač, S. Džeroski (Eds.), *Inductive Logic Programming*, Springer, Berlin, Heidelberg, 1997, pp. 1–16.
- [27] D. Singhal, H. Sirohi, A review on knowledge discovery from databases, 2022. Presented January 1, 2022.
- [28] D. Alahakoon, R. Nawaratne, Y. Xu, D. De Silva, U. Sivarajah, B. Gupta, Self-building artificial intelligence and machine learning to empower big data analytics in smart cities, *Information Systems Frontiers* 25 (2023) 221–240.
- [29] C. Poudat, H. Lungen, L. Herzberg, Investigating Wikipedia: Linguistic corpus building, exploration and analysis, John Benjamins Publishing Company, 2024.
- [30] D. Chicco, M. Warrens, G. Jurman, The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation, *PeerJ Comput. Sci.* 7 (2021) e623.
- [31] M. Alodadi, V. Janeja, Similarity in patient support forums using tf-idf and cosine similarity metrics, in: *2015 International Conference on Healthcare Informatics*, 2015, pp. 521–522.
- [32] S. Aljuneidi, W. Heuten, L. Abdenebaoui, M. Wolters, S. Boll, Why the fine, ai? the effect of explanation level on citizens' fairness perception of ai-based discretion in public administrations, in: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2024, pp. 1–18.
- [33] N. Bono Rossello, A. Simonofski, A. Castiaux, Artificial intelligence for digital citizen participation: Design principles for a collective intelligence architecture, *Government Information Quarterly* 42 (2025) 102020.
- [34] M. Borchers, T.-B. Cao, E. Bittner, Toward the ml-based analysis of citizens' inputs from e-participation in urban planning, in: *PACIS 2024 Proceedings*, 2024.
- [35] A. Solatorio, Geoformer: Predicting human mobility using generative pre-trained transformer (gpt), in: *Proceedings of the 1st International Workshop on the Human Mobility Prediction Challenge*, Association for Computing Machinery, New York, NY, USA, 2023, pp. 11–15.