

Smart Homes, Family Bonds, and Societal Resilience: A Comparative Analysis of AraBERT, MarBERT, and DistilBERT on Arabic Twitter Data

Eman Alqahtani¹, Rashid Mehmood^{2*}, Sanaa Sharaf³, Saad Alqahtany⁴

Department of Computer Science-Faculty of Computing and Information Technology,
King Abdulaziz University, Jeddah 21589, Saudi Arabia^{1,3}

Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia^{2,4}

Abstract—This study explores the concept of Smart Homes & Families by analyzing 1,174,912 Arabic tweets from Saudi Arabia to understand societal perceptions, challenges, and expectations. Recognizing that homes play a vital role in nurturing relationships, values, morals, and societal cohesion, the research emphasizes that the "smartness" of homes lies not only in technological advancements but also in supporting core family functions and contributing to sustainability. A machine learning tool was developed, integrating data collection, preprocessing, embedding generation, dimensionality reduction, clustering, visualization, and validation. The study conducts a comparative analysis of AraBERT, MarBERT, and DistilBERT (models based on Bidirectional Encoder Representations from Transformers, or BERT), identifying AraBERT as the optimal model for Arabic X (formerly Twitter) analysis. Coherence metrics and thematic evaluation were used to assess model performance. Thematic analysis revealed 22 key parameters grouped into three macro-parameters, offering a structured understanding of public discourse. The study provides policy recommendations and outlines future research directions, delivering actionable insights for stakeholders to support family well-being, societal resilience, and sustainable development through smart home technologies.

Keywords—Smart homes; smart families; sustainability; Bidirectional Encoder Representations from Transformers (BERT); AraBERT; MarBERT; DistilBERT; coherence metrics; Twitter

I. INTRODUCTION

A. Homes, Families, and Sustainability

Recent advancements in information and communication technologies (ICTs) have significantly impacted modern lifestyles, leading to the development of smart environments, cities, and societies [1], [2]. Central to this transformation are technologies like artificial intelligence (AI) and the Internet of Things (IoT), which enhance living standards by continuously monitoring surroundings and making intelligent decisions to achieve optimal outcomes. Since homes serve as the fundamental units of cities and societies, smart homes are essential for promoting smart living and are anticipated to play a key role in shaping the future of sustainable smart cities.

A smart home typically refers to a living space equipped with various network-connected devices, including remote-controlled lighting, heating systems, kitchen appliances, multimedia equipment, and electronic devices, often integrated with sensors that produce and process large amounts of data [3]–

[6]. This data is continuously analysed using AI, big data analytics, and large-scale distributed computing to provide real-time insights for enhancing comfort, security, efficiency, and sustainability. Recent advances in fog and edge computing have further optimized smart home operations by reducing response time delays that were previously associated with cloud-based processing [7], [8]. However, despite these technological advancements, current academic literature and commercial developments have predominantly focused on functional aspects of smart homes, such as ambiance management [9], energy optimization [4], security management [10], appliance control [11], and healthcare services [3]. While these functions are undoubtedly important, homes represent far more than just physical spaces or environments for convenience and security. They are complex social constructs that embody security and control, activity, relationships and continuity, and identity and values.

More importantly, homes play a central role in nurturing responsible citizens who contribute positively to society and help prevent harm to it. They serve as the foundation for raising future leaders, innovators, entrepreneurs, educators, researchers, policymakers, healthcare providers, artists, scientists, and changemakers, individuals who will collectively strengthen the triple bottom line of social, environmental, and economic sustainability. Homes are where ethical values are instilled, good behaviours are promoted, harmful behaviours are prevented, and critical thinking, creativity, and resilience are cultivated. They provide the environment for education and lifelong learning, emotional growth, and cultural continuity, all of which are essential for building cohesive and prosperous societies.

In recent years, a global decline in marriage rates, increasing divorce rates, and a weakening of interpersonal relationships have raised concerns about the long-term sustainability of societies, posing risks not only to social cohesion but also to the continuity of human existence. The erosion of family structures threatens to accelerate social fragmentation, potentially leading to broader societal instability. Additionally, the decline in shared morals and values within family units further undermines the foundations of cohesive societies. True homes, therefore, must be environments that strengthen human relationships, nurture moral and ethical values, and support social bonds that sustain future generations. They should foster meaningful connections that encourage family formation, promote child-rearing in stable

environments, and cultivate a sense of shared responsibility for the well-being of communities and societies.

A truly smart home should therefore not be defined merely by technological sophistication but by how effectively technology supports these core societal and familial objectives. The smartness of a home lies in its ability to strengthen relationships, promote positive behaviours, and enable families to function as cohesive units that contribute to a sustainable society. Technologies should facilitate education, foster innovation, and help families adapt to evolving social challenges rather than merely providing entertainment or operational convenience. Smart homes should thus be seen as enablers of societal well-being, where technology plays a supporting role in achieving the primary goals of homes and families.

Relatedly, with the rise of social media platforms such as X (formerly Twitter), there is an unprecedented opportunity to analyse public discourse and sentiments related to families and homes. Arabic Twitter data is especially valuable as it offers rich insights into how families in Arabic-speaking societies discuss their living environments, societal challenges, and technological expectations. However, extracting meaningful patterns from such unstructured data requires advanced natural language processing (NLP) techniques specifically designed for Arabic text analysis. While existing research has explored smart home technologies from the perspective of automation and efficiency [6], [12]–[17], little attention has been given to their intersection with family dynamics, social cohesion, and cultural values.

This gap highlights the need for a comprehensive analysis that bridges technological advancements with family dynamics, using context-sensitive data and advanced analytical models.

B. Our Approach and Contributions

Building upon the conceptual understanding that smart homes should not be limited to technological convenience, entertainment, or security, this study explores how public discourse reflects the essential societal roles of homes and families. The primary aim of this study is to analyse public discourse to uncover the key functions, roles, and challenges associated with smart homes and families. The research focuses on identifying the core societal priorities related to homes and families and examining the barriers that hinder them from fulfilling these roles. By analysing Arabic Twitter data, this study provides insights into how homes contribute to relationship building, ethical development, and long-term societal resilience.

The findings are intended to guide policymakers, technology developers, and other stakeholders in designing smart home technologies and solutions that genuinely support family well-being and societal sustainability. In essence, this research frames the concept of "smartness" in homes as a reflection of how effectively technology supports the real objectives of homes and families, rather than being defined solely by the accumulation of advanced technological features. By grounding technological advancements in these core societal functions, the study contributes to creating smart living environments that strengthen the social fabric and promote sustainable societal progress.

To achieve its objectives, this study employs a data-driven approach to explore the key functions, roles, and challenges associated with smart homes and families by analysing 1,174,912 Arabic tweets. The Twitter platform was selected as the primary data source because it offers rich, real-time insights into societal values, aspirations, and challenges, reflecting the public discourse surrounding homes and families. The data collection process was carefully designed to ensure that the dataset captured relevant discussions, focusing on the core societal objectives of homes and families. This careful curation ensured that the collected data provided comprehensive coverage of public conversations, allowing for meaningful analysis aligned with the research aims.

A central methodological component of this study is the comparative evaluation of three advanced Bidirectional Encoder Representations from Transformers (BERT) models: AraBERT, MarBERT, and DistilBERT. The performance of these models was evaluated using quantitative coherence metrics, which measure how semantically coherent and interpretable the generated clusters and parameters are. The results showed that the AraBERT model outperformed the others, providing the most coherent thematic structures and interpretable clusters within the Arabic Twitter data. In addition to the quantitative evaluation, this research also conducted a subjective assessment of the thematic boundaries of the clusters generated by each model. This involved a qualitative review of the clusters and parameters, focusing on their clarity, distinctiveness, and relevance to the overarching themes related to homes and families. The subjective evaluation confirmed the better performance of the AraBERT model, which demonstrated a greater ability to define clear thematic structures that aligned closely with the study's objectives. This dual approach, combining quantitative and qualitative evaluations, provided a robust validation of the selected model, ensuring the reliability of the findings.

To ensure data quality, the study implemented a refined data preprocessing pipeline. This process involved systematically filtering out irrelevant content, including promotional material, advertisements, and home-related services that were unrelated to the core themes of interest. The preprocessing step was critical in ensuring that the final dataset reflected focused discussions on the fundamental objectives of homes and families, thereby enhancing the accuracy and relevance of the subsequent analysis.

A detailed thematic analysis was conducted, resulting in the identification of 22 key parameters, which were grouped into three macro-parameters: Nurturing Families, Education & Career Development, and Family Challenges. These parameters provide a structured understanding of how the public perceives the societal roles of homes and families, highlighting societal priorities, challenges, and opportunities related to smart living environments. The analysis offers deep insights into how homes can contribute to relationship building, ethical development, and societal resilience. These insights are further complemented by practical policy recommendations, designed to guide policymakers, technology developers, and other stakeholders in designing smart home solutions that genuinely support family well-being and societal sustainability. To summarize, the key contributions of this research are as follows.

1) Developed a machine learning tool integrating key components such as data preprocessing, embedding generation using AraBERT [18], MarBERT [18], and DistilBERT [19], dimensionality reduction with UMAP [20], clustering via HDBSCAN [21], and topic extraction using class-based TF-IDF (c-TF-IDF) [22]. The tool also includes visualization techniques using Matplotlib [23], Seaborn [24], and Plotly [25], along with validation processes involving internal and external evaluations to ensure the robustness and reliability of the findings.

2) Curated a large-scale dataset comprising 1,174,912 Arabic tweets, collected using the Twitter API v2 with geolocation filtering for Saudi Arabia.

3) Conducted a comparative analysis of AraBERT, MarBERT, and DistilBERT, utilizing our machine learning tool based on quantitative coherence metrics and subjective thematic evaluations.

4) Delivered a focused thematic analysis, identifying 22 key parameters grouped into three macro-parameters, and provided an information structure (taxonomy) of smart homes and families, offering a structured understanding of societal perceptions related to the topic.

5) Provided practical policy recommendations and outlined future research directions, offering actionable insights for stakeholders to support family well-being, societal resilience, and sustainable development through smart home technologies.

This study builds upon our previous research on Smart Homes & Families [26], which combined Scopus academic literature to explore academic perspectives and Twitter data to capture public sentiment. While the earlier study provided foundational insights into the intersection of smart homes and family dynamics, it had certain limitations that this research addresses.

First, we improved the previously developed software tool, enhancing its capabilities in data preprocessing, embedding generation, clustering, and visualization, which enabled a more robust and efficient analysis. Second, we refined the data collection process by using a targeted set of search keywords, ensuring the dataset captured more relevant discussions on the core objectives of homes and families. Following data collection, we enhanced the preprocessing pipeline by removing irrelevant tweets, including home-related promotional content, resulting in a cleaner dataset and more focused insights. Third, unlike the previous study, which utilized a single BERT model, this research conducts a comparative analysis of AraBERT, MarBERT, and DistilBERT to evaluate their performance. This comparison led to improved parameter discovery, deeper data analysis, and the extraction of meaningful insights, with AraBERT identified as the optimal model for Arabic Twitter analysis. Additionally, this study delivers a more detailed thematic analysis, providing comprehensive findings, policy recommendations, and future research directions that were not extensively covered in the earlier work. These advancements represent a significant improvement over the previous study, delivering deeper insights, greater methodological robustness, and practical recommendations for the development of smart

home technologies that support family well-being and societal sustainability.

The remainder of this paper is organized as follows. Section II reviews related literature in the areas of homes and families, smart home technologies, and Twitter-based analytics. Section III presents the methodology and tool architecture used for data processing and parameter discovery. Section IV describes the three BERT models and their performance comparison. Section V details the results of the thematic analysis using AraBERT. Section VI discusses the key findings, offers policy recommendations, and outlines directions for future research. Finally, Section VII concludes the study.

II. RELATED WORK

This review establishes the research gap addressed by this study by summarizing key works in three areas: meanings and concepts of homes, technological aspects of smart homes, and social media analytics using Twitter data. While extensive research exists, no previous work directly aligns with the specific focus of this study.

The concept of home has been examined from various perspectives. Gram-Hanssen and Darby [27] highlighted discrepancies between technical research on smart homes—focusing on IoT, AI, and automation—and broader conceptual meanings related to relationships, values, identity, and security. They grouped ten meanings of homes from Després [28] into four categories: security and control, activity, relationships and continuity, and identity and values, although their primary focus remained on energy management. Mitty and Flores [29] defined home in terms of physical space, geography, and relationship-building, while other studies examined how age, culture, and health conditions influence perceptions of home, such as Hatcher et al. [30] on older adults and Lewin [31] on elderly immigrants.

In the context of smart homes, research has primarily focused on technological aspects. Reviews such as Marikyan et al. [13] emphasized the need to consider the user perspective for better adoption of smart home technologies. DeFranco and Kassaba [14] proposed a taxonomy for smart home research but noted the lack of consensus on definitions and research directions. Pira [15] identified trust, service satisfaction, reliability, and privacy as key social barriers to adoption. Additionally, Li et al. [6] summarized core research themes in smart homes, including AI for home automation, energy management, and home-based healthcare. Other studies, such as Choi et al. [12], explored smart IoT (SHIoT) dimensions like household automation, network security, and energy efficiency, while Singh et al. [16] focused on IoHT for elderly health monitoring. Li et al. [17] further highlighted the role of IoT, cloud computing, and machine learning in balancing energy efficiency and user comfort.

The use of Twitter data in research has gained prominence due to its richness and immediacy. For instance, Alotaibi [32] introduced Sehaa, a big data analytics tool for healthcare in Saudi Arabia. Alomari [33] developed Iktishaf, leveraging Twitter data to detect traffic-related events. Studies such as Sauris et al. [34] analyzed security concerns in smart living environments, while numerous works explored COVID-19-

related issues using Twitter analytics (Su et al. [35], Abdulaziz et al. [36]). Additionally, Alswedani et al. [37] examined governance parameters in the education sector using Twitter-based analytics. Mental health was explored using Twitter data in [38], highlighting a critical dimension of well-being that deeply influences family dynamics, home environments, and broader societal stability.

Our previous work [26] also contributed to this area by analysing Twitter data alongside academic literature to explore the concept of Smart Homes & Families. However, that study primarily focused on general discussions and utilized a single BERT model, offering initial insights. The current research advances this earlier work by employing a comparative analysis of multiple BERT models, improving data preprocessing techniques, and delivering deeper thematic insights along with practical policy recommendations.

The literature establishes that, while extensive work exists across the domains of home concepts, smart home technologies, and Twitter-based analytics, no prior study directly investigates the intersection of smart homes and families through Arabic Twitter data using a comparative BERT-based approach, while also improving data preprocessing techniques, delivering deeper thematic insights, and providing practical policy recommendations, as achieved in this research.

III. METHODOLOGY

This section presents our methodology and tool design for analyzing and identifying parameters from an Arabic Twitter dataset concerning Smart Homes & Families. The tool architecture is depicted in Fig. 1.

Our data collection process utilized the Twitter platform as the primary source to capture public opinions related to Smart Homes & Families. Using the Twitter API v2 from January to June 2022, we collected 1,174,912 Arabic tweets. Table I provides a complete list of search query terms and their English translations. We applied geolocation filtering to extract tweets from Saudi Arabia. The tweets were retrieved in JSON format with attributes such as 'created_at,' 'text,' 'geo,' and 'place,' which were later extracted and stored in a CSV file.

The preprocessing process thoroughly cleaned and prepared the dataset for analysis. The collected tweets were loaded into a Pandas DataFrame, where duplicate entries and non-Arabic language tweets with similar scripts (e.g., Urdu, Persian, Central Kurdish) were removed based on the "lang" attribute to maintain the dataset's relevance. Irrelevant characters, including English letters, numbers, punctuation, hashtags, mentions, emails, emojis, links, and extra spaces, were systematically eliminated to ensure the cleanliness of the data. Additionally, promotional and advertisement tweets were removed. Arabic diacritics, which include short vowels, nunation, and shadda diacritics, were removed also to normalize the text. Words containing different forms of Alif (أ, إ, ؤ), Taa Marbutah (ة), and Yaa (ي) were standardized to their basic forms for instance, Alif was replaced by bare Alif (ا), Taa Marbutah by haa (ه), and Yaa by dotless Yaa (ى). Following normalization, the tweets were tokenized and saved in a CSV file. Notably, stop words were retained during preprocessing to preserve the full context necessary for accurate embeddings, allowing their removal post-

embedding generation using the Count Vectorizer component of the BERT model. Through these preprocessing steps, we retained only relevant and well-organized data for the next process.

TABLE I. KEY ARABIC VOCABULARY WITH ENGLISH TRANSLATION USED IN THE TWITTER DATA COLLECTION

Arabic	English	Arabic	English	Arabic	English	Arabic	English
القيم	Values	تنشئة	Nurturing	أخت	Sister	والدة	Mother
الأخلاق	Moral	أم	Mother	صداقة	Companionship	الوالدين	Parent
التربية	Nurturing	أب	Father	طفل	Baby	الأخوة	Brothers
الأبناء	Children	أخ	Brother	والد	Father	الأخوات	Sisters

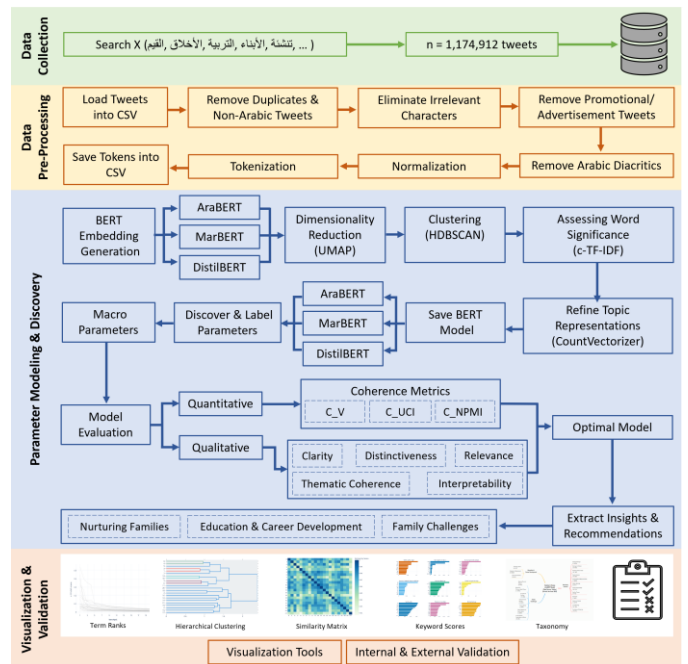


Fig. 1. Smart homes and families: Methodology and design.

For parameter modeling and discovery, we employed three distinct BERT-based models AraBERT, MarBERT, and DistilBERT. Each model facilitated the extraction of contextual relationships between words by converting the pre-processed tweets into dense numerical representations. To effectively manage the high dimensionality of these embeddings, we utilized the Uniform Manifold Approximation and Projection (UMAP) algorithm, which preserved both global and local structures essential for meaningful clustering. Subsequently, we applied the HDBSCAN algorithm to group the reduced embeddings into clusters based on semantic similarities, optimizing key parameters such as min_cluster_size and min_samples to improve clustering quality.

The significance of words within each identified cluster was assessed using class-based TF-IDF (c-TF-IDF) scores, which measure word importance by comparing term frequency within a cluster to its overall occurrence across the corpus. This approach enabled the derivation of keyword-based descriptions for each cluster. Additionally, we integrated CountVectorizer

with c-TF-IDF to enhance topic representations by eliminating stop words, thereby refining the quality of the identified topics.

An iterative fine-tuning process was applied to adjust the `nr_topics` parameter in each model, determining the final number of thematic clusters. These clusters underwent further refinement to ensure coherence and relevance, which involved removing irrelevant clusters, merging thematically similar ones, and assigning appropriate labels based on domain expertise. The labeled clusters (referred to as parameters) were subsequently grouped into broader macro-parameters through the use of similarity matrices, hierarchical clustering, and quantitative analysis, enhancing interpretability and offering a structured perspective on Smart Homes and Families research.

The validation process included both internal and external evaluations. Internal validation assessed the relevance of each tweet to its corresponding cluster, ensuring meaningful relationships between the texts and clusters. External validation involved comparing the derived parameters with established research findings, thereby reinforcing the credibility of the results. Visualization techniques, such as term ranking plots, hierarchical clustering dendrograms, and similarity matrices, were employed to interpret and present the findings. These visualizations were generated using Python libraries, including Matplotlib [23], Seaborn [24], and Plotly [25], facilitating a detailed analysis of the dataset and topic structures.

This methodological framework enabled the extraction, processing, and analysis of a large volume of Arabic Twitter data, leading to the identification of key parameters and macro-parameters related to families and homes in Saudi Arabia. The combination of rigorous preprocessing, advanced modeling techniques, and thorough validation ensured the robustness and reliability of the findings, providing a strong foundation for future research on Smart Homes and Families.

IV. MODEL COMPARISON

This section describes the three models applied in our analysis, AraBERT, MarBERT, and DistilBERT, and compares their performance using coherence metrics.

A. AraBERT Model

AraBERT is a language model designed specifically for Arabic, pretrained on a dataset comprising Modern Standard Arabic news content from a variety of Arabic media sources. The initial version, AraBERTv0.1, includes 77 million sentences and 2.7 billion tokens, corresponding to approximately 23 gigabytes of text. The second version expanded the pretraining data by 3.5 times, resulting in a total of 77 gigabytes of text. Structurally, AraBERT consists of 12 transformer layers, each containing 768 hidden units and 12 self-attention heads, totaling 110 million trainable parameters. To enhance its performance with dialectal Arabic, the model was fine-tuned on 12,000 sentences covering various Arabic dialects [39]. We used the second version (AraBERTv0.2-Twitter-base), which is fine-tuned on Arabic dialects and Twitter data.

B. MarBERT Model

MarBERT is an Arabic language model designed to handle both Modern Standard Arabic (MSA) and various Arabic dialects. It was trained on a large corpus of Twitter data,

encompassing one billion tweets and nearly 128 gigabytes of text, with approximately 15.6 billion tokens in total. The architecture of MarBERT includes 12 transformer layers, each containing 768 hidden units and 12 self-attention heads, resulting in around 160 million trainable parameters. Notably, the model performs effectively without the next-sentence prediction (NSP) component, which was deliberately excluded by the developers due to the short length of tweets [39].

C. DistilBERT Model

DistilBERT is a distilled version of the original BERT model, designed to be lighter and faster while retaining much of BERT's capabilities. It supports multiple languages, including Arabic. Unlike BERT, which features a more extensive architecture, DistilBERT has a reduced structure with 6 transformer layers, 12 attention heads, and a hidden size of 768 dimensions, resulting in approximately 66 million parameters. DistilBERT is trained using a distillation process to mimic BERT's behavior, focusing on speed and size reduction. It achieves about 97% of BERT's performance on various benchmarks while being 60% faster and occupying less disk space, making it suitable for resource-constrained environments [19].

D. Model Comparison and Evaluation

While AraBERT and MarBERT are specifically designed for Arabic text, DistilBERT is a multilingual model with Arabic support. To evaluate these models, coherence metrics were used to assess their performance in topic modeling. Coherence values are essential for determining how well a model semantically integrates the top-scoring words within topics, helping to distinguish meaningful topics from those formed by statistical artifacts [40]. Three metrics were employed for this purpose. The `C_V` metric evaluates semantic similarity among words using external word embeddings, aligning well with human judgment by emphasizing topic interpretability. The `C_UCI` metric measures word co-occurrence, focusing on how well topics capture significant patterns in the data. Finally, the `C_NPMI` metric assesses the strength of associations between words, balancing statistical significance and interpretability by normalizing mutual information to account for word prevalence. These metrics were implemented using the Gensim library [41].

Fig. 2, 3, and 4 depict the coherence scores (`C_V`, `C_UCI`, and `C_NPMI`, respectively) for AraBERT, MarBERT, and DistilBERT, respectively, with the x-axis showing categories of Top Words (ranging from Top 5 to 20) and the y-axis representing coherence values. We note that, in each category, AraBERT consistently achieved higher coherence scores, demonstrating its robust performance in generating semantically coherent and interpretable topics compared to MarBERT and DistilBERT. This highlights AraBERT's strength in capturing meaningful patterns within Arabic text data.

The coherence metrics in the figures show how well the language models (AraBERT, MarBERT, DistilBERT) capture semantic relationships between top words. A positive coherence value indicates that the model identifies meaningful connections, demonstrating a good understanding of the text. In contrast, a negative value suggests difficulty in capturing semantic relationships. The AraBERT model was chosen because it achieved better coherence scores across the three

metrics (C_V, C_UCI, and C_NPMI) than MarBERT and DistilBERT, indicating its superior ability to capture semantic relationships.

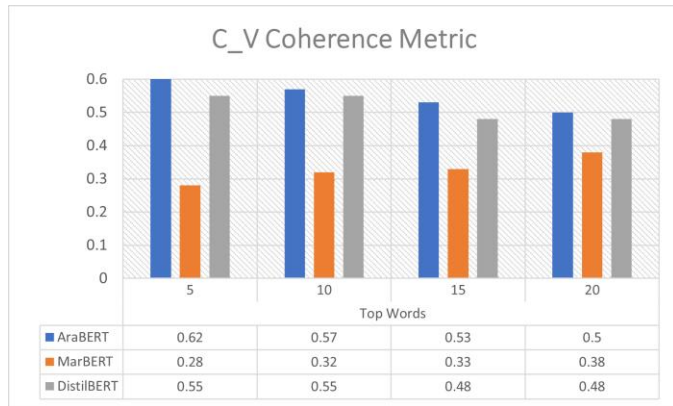


Fig. 2. C_V coherence metric for the three models with varying numbers of top words.

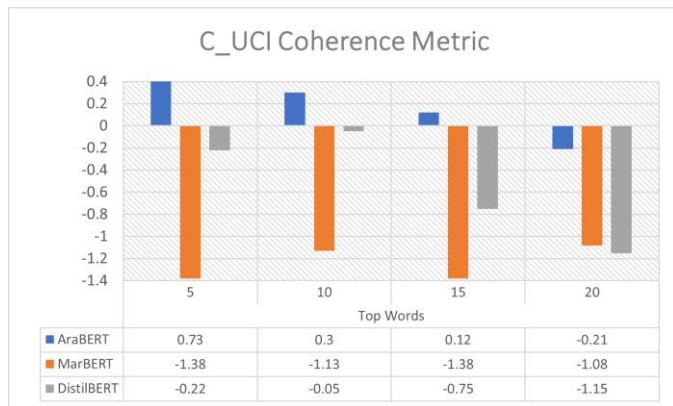


Fig. 3. C_UCI coherence metric for the three models with varying numbers of top words.

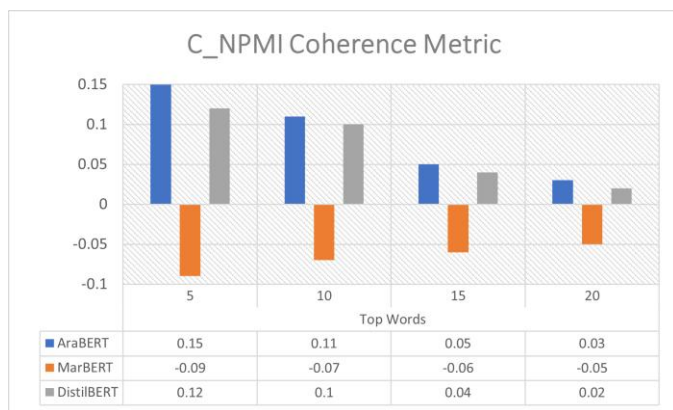


Fig. 4. C_NPMI coherence metric for the three models with varying numbers of top words.

Table II presents the average coherence scores for AraBERT, MarBERT, and DistilBERT across three metrics (C_V, C_UCI, and C_NPMI), evaluating their effectiveness in generating semantically coherent topics. AraBERT consistently achieves the highest scores, demonstrating its strong ability to produce coherent topics. DistilBERT performs moderately well

but shows lower coherence in the C_UCI metric, where it records a negative score. MarBERT has lower scores across all metrics, particularly C_UCI, indicating limitations in capturing meaningful topic patterns. This comparison highlights the models' relative strengths in maintaining topic coherence, with AraBERT achieving the best performance.

TABLE II. THE AVERAGE COHERENCE SCORES FOR THE THREE MODELS ACROSS THREE METRICS

Metric	AraBERT	MarBERT	DistilBERT
C_V	0.555	0.328	0.515
C_UCI	0.235	-1.243	-0.543
C_NPMI	0.085	-0.068	0.070

Moreover, our qualitative evaluation assessed the clarity of thematic boundaries within the clusters, the depth of topic categorization, and the resulting taxonomies comprising parameters and macro-parameters. Our analysis revealed that AraBERT not only maintains well-defined thematic structures but also captures a broader spectrum of socio-political, spiritual, and family-related distinctions compared to MarBERT and DistilBERT. Unlike the other models, AraBERT effectively balances parameter-level family dynamics, such as traditional roles and responsibilities, with macro-level societal influences, including global crises, maternal and child health, and spiritual discourse. This holistic representation makes AraBERT a better and coherent model for analysing smart homes and families. Given its ability to integrate both personal and systemic factors, we selected AraBERT as the primary model for our analysis, and the remainder of this paper focuses on the insights derived from its results.

V. RESULTS AND ANALYSIS (ARABERT MODEL)

This section outlines the parameters identified by the AraBERT model from the Arabic Twitter dataset, representing public perceptions of Smart Homes and Families in Saudi Arabia. A total of 22 parameters were detected and subsequently grouped into three macro-parameters. Section A presents the quantitative analysis and the taxonomy of these parameters and macro-parameters, while Sections B to D detail each macro-parameter.

A. Quantitative Analysis

The AraBERT model identified 22 thematic clusters from the dataset. Using a combination of domain expertise, similarity matrices, hierarchical clustering, and other quantitative techniques, these clusters were assigned descriptive labels. The parameters were then categorized into three macro-parameters: Nurturing Families, Education and Career Development, and Family Challenges. Fig. 5 illustrates the Smart Homes & Families taxonomy, showing the macro-parameters as the primary categories on the first level of branches, and the parameters as subcategories on the second level, along with their corresponding cluster numbers and tweet counts. For example, the parameter "Parents' Roles and Responsibilities (1, 54,635)" corresponds to Cluster 1, which contains 54,635 tweets. This taxonomy provides a structured way of categorizing information related to the Smart Homes & Families domain.

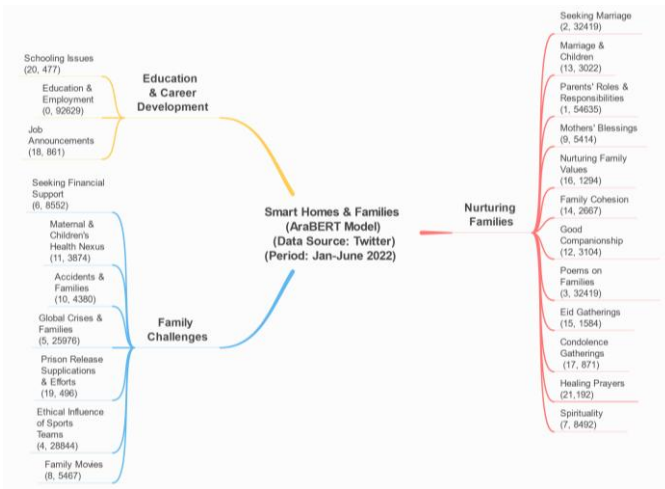


Fig. 5. Smart homes and families: A Taxonomy.

For quantitative analysis, various visualization tools were employed to evaluate the extracted clusters and parameters. These tools included term ranks, hierarchical clustering, similarity matrices, and keyword scores. While each cluster is associated with specific keywords, not all these keywords effectively represent their respective clusters. Fig. 6 illustrates the number of keywords required to accurately describe a parameter. On average, only the top seven to ten terms provide a meaningful description for each parameter.

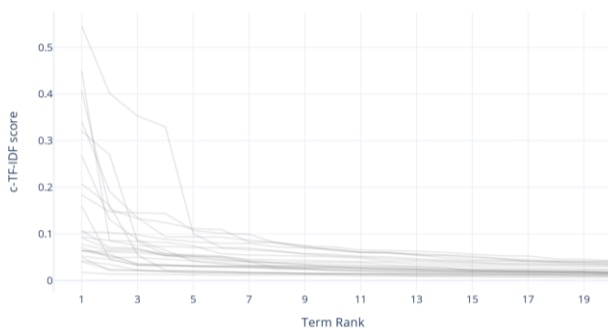


Fig. 6. Cluster term ranks.

Fig. 7 shows the hierarchical clustering of 22 clusters within Smart Homes and Families. Clusters 12, 14, and 16, due to their high similarity, have been grouped together in a macro-parameter. The similarity matrix in Fig. 8 depicts the relationships between different clusters in Smart Homes & Families. The darker blue cells indicate higher similarity scores, while the lighter green cells represent lower similarities. For instance, the dark blue cell at the intersection of Cluster 1 (Parents' Role & Responsibilities) and Cluster 14 (Family Cohesion) suggests these two clusters share common features and are closely related. This type of visualization helps to highlight the conceptual connections between the various themes and topics explored within this research field, allowing us to discover information structure within the field of Smart Homes & Families through parameter discovery and refinement.

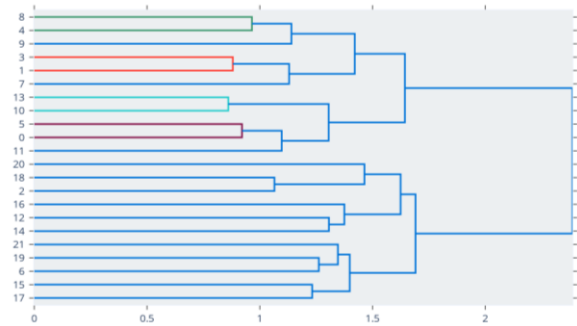


Fig. 7. Hierarchical clustering diagram.

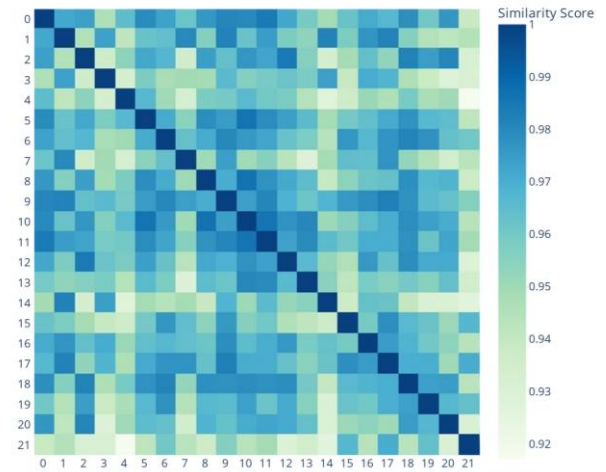


Fig. 8. Cluster similarity matrix.

Later, in the respective subsections for each macro-parameter (Sections B to D), we will present visualizations of the top 10 keywords associated with their corresponding parameters. These keywords are ranked based on their c-TF-IDF importance scores. The visualizations feature horizontal lines representing the magnitude of the c-TF-IDF scores, while vertical lines display the corresponding parameter keywords.

B. Nurturing Families

Nurturing Families encompasses 12 distinct parameters (see Fig. 9) that collectively represent the values, traditions, and responsibilities that shape strong and loving family bonds. It highlights the roles of parents, the importance of companionship, and the celebration of familial unity across significant life moments.

The journey of building and nurturing a family begins with Seeking Marriage, representing the significant step of finding a life partner to begin a family. The next phase, Marriage and Children, focuses on the joys and responsibilities of building a life together and raising the next generation. Within the family, Parents' Roles and Responsibilities play a critical part in providing guidance and creating a nurturing environment, with Mothers' Blessings celebrating the irreplaceable role of mothers. Nurturing Family Values emphasizes instilling ethics and traditions to foster growth, while Family Cohesion

underscores the importance of teamwork, harmony, and mutual respect. The role of Good Companionship extends to fostering supportive relationships both within and outside the family, creating a sense of belonging. Emotional expressions of love and gratitude are captured in Poems on Families, while Eid Gatherings highlight festive moments that reinforce familial bonds and traditions. During times of loss, Condolence Gatherings emphasize family solidarity and support, while Healing Prayers reflect the spiritual comfort sought in moments of hardship, fostering hope and resilience. Lastly, Spirituality captures the various aspects of family spirituality, including religion, faith, and religious education, which play a vital role in guiding the family, reinforcing its values, and building resilience. Overall, Nurturing Families captures the essence of fostering love, respect, and growth within the family as a cornerstone of life and society.



Fig. 9. c-TF-IDF Keyword Scores (Nurturing Families).

These ideas resonate in the tweets of the parameters. For Seeking Marriage, tweets highlight the preferences and aspirations of individuals seeking a life partner, as captured in, “I work in a government job, live in my own house, and seek marriage”, and “A single man, committed to prayer, educated, and from [X] city looking for marriage”. Similarly, Marriage & Children focuses on the joys and responsibilities of marital life and raising children, as expressed in the tweets, “Children are the most precious thing to me I would do anything to protect and nurture them”, “You think raising kids is easy? They are a lifelong responsibility...”, and “I love children and their innocence they bring so much joy”. For Parents' Roles & Responsibilities, tweets reflect the significant role parents play in shaping their children's character and addressing challenges, as seen in, “Children need role models, not critics. Spend time with them if you don't make time for them, they won't have a place for you in their lives”, “The problem of bullying often starts at home, within the family itself, as parents bear a great responsibility toward their children ...”, and “Throughout history, women have been the primary shapers of generations. This is why Islam has emphasized their role as nurturers of children and cultivators of divine values”.

Mothers' Blessings emphasizes the irreplaceable role of mothers, beautifully captured in the tweets, “A mother's voice in the house is the greatest blessing...”, “A mother remains the haven in every stage of life ...”, and “The only person who truly deserves all your love...”. For Nurturing Family Values, the focus on instilling ethics and morals is evident, as seen in, “Respect is a sign of good upbringing, not weakness, and that

apologizing is a virtue, not humiliation”, “Teach your children that a person's true beauty lies in their morals, and that appearance is not everything”, and “Beauty isn't just something you see, but something you reflect through your actions and character”. Similarly, Family Cohesion highlights the bonds that keep families united, reflected in the tweets, “Siblings are the flowers nurtured by a mother's love, growing stronger together”, “A father is one of the greatest blessings, not just as a guide but as a pillar that keeps the family strong and united”, and “A mother is the heart of the family, bringing everyone together through love and care”.

Good Companionship explores the importance of meaningful relationships, as one tweet notes, “Choose friends who guide you toward good...”, while another describes friendships as blessings, “Mothers, siblings, and friends are blessings...”, and others emphasize the comfort of unique connections, “There's a bond called friendship, and it's beautiful...”. In Poems on Families, the beauty of familial love and connection is highlighted in the tweets, “Be kind to people and maintain good manners...”, and “Traveling, work, and dust no matter how tired I am, I'll always return to the home...”.

Eid Gatherings capture the joy of festive traditions, as seen in the tweets, “It's part of the religion to show joy on Eid...”, “Our family's house unites us with love on every Eid”, and “Enjoy Eid with family and friends”. For Condolence Gatherings, tweets reflect on moments of loss and solidarity, as expressed in, “To God we belong, and to Him, we shall return. Today we lost our dear friend ...”, “Heartfelt condolences to the [X] family on the passing of [Name]. May their soul rest in peace...”, and “We ask God to grant her mercy, forgive her, and comfort...”. Similarly, Healing Prayers focus on the spiritual comfort sought in times of illness, as shared in, “I ask God, the Lord of the Great Throne, to heal...”, “May God heal him, grant him wellness, and bring joy...”, and “I pray to God during these holy days to heal my father...”. Lastly, Spirituality emphasizes faith and religious practices as a cornerstone of family values and resilience, as evident in the tweets, “Tomorrow is Monday fasting. Blessed is the one who has intended it and devoted their effort sincerely”, and “The one who guides others to goodness is like a beacon of light, sharing in the reward of their deeds”.

Supporting parental roles and promoting family cohesion are vital for fostering strong familial foundations. Governments and organizations can collaborate on parenting workshops that teach communication, emotional intelligence, and responsibility. Digital platforms designed to modern parenting challenges would further enhance these efforts. Public campaigns emphasizing the value of marriage and premarital counseling programs could also contribute to societal stability. Additionally, cultural and spiritual initiatives, such as family gatherings during events such as Eid and shared prayer sessions, can strengthen intergenerational bonds. Mental health and spirituality services and community-led condolence sessions should be accessible to families facing grief, providing emotional support during challenging times. Programs that celebrate family through arts and literature, alongside youth engagement initiatives, can further reinforce these values by inspiring communities and fostering connections between generations.

Future work should focus on studying evolving family dynamics and the impact of societal changes on relationships. Research into the effects of digitalization, shifting parental roles, and family interactions will provide valuable insights. Community-based interventions, such as pilot programs, can measure the impact of family-centric policies, while the role of spirituality in healing and intergenerational cohesion should be further explored. Technological solutions, such as AI-driven applications and social media analytics, can support families and uncover public sentiment about family togetherness. Longitudinal studies can evaluate how family-centric upbringing influences career and personal life outcomes, emphasizing the role of education in reinforcing family values. Together, these recommendations aim to enhance family well-being and address emerging challenges in maintaining strong familial bonds.

C. Education and Career Development

Education and Career Development (see Fig. 10) highlights the interconnected roles of schooling, education, and professional opportunities in shaping individual and societal progress. It emphasizes the importance of creating supportive and secure learning environments, bridging education with employment, and promoting equitable career opportunities. Together, these dimensions aim to empower individuals and families while fostering economic and societal growth.



Fig. 10. c-TF-IDF Keyword Scores (Education and Career Development).

Schooling Issues focuses on foundational education, addressing challenges and themes related to early childhood, primary, and secondary schooling. Keywords such as "الابتدائية" (elementary), "الاطفال" (children), and "التعليم" (education) highlight the importance of nurturing young learners in secure and well-structured environments. Discussions about "العودة" (return to school) and "رياضة الاطفال" (kindergarten) reflect efforts to create seamless transitions for students in dynamic or disruptive circumstances (the Covid-19 pandemic). The mention of "نظام نور" (Noor system) emphasizes the role of technology and systemic improvements in streamlining educational processes while ensuring "امنه" (safe) learning conditions for all students. This dynamic is exemplified in various discussions highlighted in the tweets. For instance, the role of families in supporting children's education is frequently emphasized, as captured in tweets such as, "الانضباط المدرسي وعدم الغياب يبدأ من... الاسره" ("School discipline and attendance start with the family"), underscoring the importance of parental involvement in fostering consistency and commitment to education. Similarly, the importance of motivation in improving academic achievement is reflected in tweets such as, "دور الاسره في تنميته... الدافعيه" ("The role of the family in developing motivation..."), highlighting how familial encouragement can enhance student performance. Additionally, the accessibility of educational

systems is exemplified in the tweet, "بدء تسجيل الاطفال في الروضه" ("Registration of children in public kindergartens through the Noor system"), showcasing the integration of technology in streamlining school enrollment processes.

Education and Employment bridges the gap between learning and professional opportunities, focusing on the role of education in preparing individuals for work and societal contributions. With keywords such as "التربية والتعليم" (education and learning), "العمل" (work), and "الاسره" (family), this parameter highlights the balance between home life and professional aspirations. It also emphasizes the importance of societal support, as seen in mentions of "وزارة التربية" (Ministry of Education), which plays a vital role in shaping policies that align education with workforce demands. Furthermore, terms such as "القيم" (values) and "المنزل" (home) underscore the integration of cultural values into learning and employment frameworks, promoting an inclusive and balanced approach to societal growth. This collaborative effort is evident in tweets such as, "...على الاسرة والمعلمين والمعلمات والهيئة الإدارية تحفيز الطلاب" ("Families, teachers, and administrative staff must motivate students"), highlighting the collective responsibility to ensure a strong educational foundation. Success stories in educational innovation are also celebrated, as reflected in a tweet, "حقق التعليم" ("Education achieved significant success through the Madrasati platform"), emphasizing the role of digital platforms in revolutionizing education. Furthermore, institutional involvement is evident in the tweet, "وزارة التربية" ("The Ministry of Education publishes instructions and guidelines"), which underscores the role of government bodies in preparing students for critical academic milestones.

Job Announcements captures discussions about employment opportunities, job-seeking platforms, and career advancement. The keywords "وظائف" (jobs), "فرص" (opportunities), and "رواتب" (salaries) reflect a strong focus on providing access to diverse career paths and ensuring fair compensation. The frequent mention of digital platforms such as "واتساب" (WhatsApp) and "تليجرام" (Telegram) underscores the role of technology of social media in democratizing access to job markets and connecting job seekers with employers. Discussions also span industries and demographics, highlighted by terms such as "رجال" (men) and "نساء" (women), emphasizing inclusivity and diversity in employment opportunities. For example, a tweet states, "تعلن وزارة الموارد البشرية عن توفر وظيفه" ("The Ministry of Human Resources announces a remote job vacancy"), highlighting efforts to make employment accessible across various qualifications and demographics. Similarly, private sector initiatives are evident in various tweets reflecting the active role of companies in creating diverse career opportunities for both genders through dedicated employment events.

Together, these discussions illustrate the seamless integration of schooling, education, and career pathways within Education and Career Development, demonstrating the collaborative efforts of families, institutions, and industries in fostering personal and professional growth while contributing to societal advancement.

To strengthen the integration of digital platforms in education, governments should expand the reach of tools such as Madrasati, ensuring equal access for students in remote or underserved areas. Training programs for teachers and administrative staff should be prioritized to maximize the effectiveness of these platforms in fostering academic and developmental progress. Policies should also focus on promoting holistic student development by encouraging collaboration among families, educators, and administrators, supported by awareness campaigns that emphasize parental engagement in education. Furthermore, aligning educational curricula with labor market demands, especially in technology and sustainable development, will better prepare students for future workforce needs. Vocational training programs targeting younger students should be introduced to promote career readiness and skill-building from an early age.

On the employment front, policies should support remote work and freelancing by introducing fair regulations and fostering partnerships between the public and private sectors. Initiatives such as NEOM's recruitment forum and announcements from the Ministry of Human Resources highlight the importance of public-private collaboration in job creation. Promoting inclusive hiring practices, particularly for women and marginalized groups, is essential to fostering a diverse and equitable workforce. Additionally, governments should invest in feedback mechanisms to assess the effectiveness of implemented policies, supported by research on the long-term impact of digital learning and employment programs. Future efforts should focus on leveraging AI and global collaboration to adopt cutting-edge practices, ensuring that education and employment systems remain adaptive, inclusive, and aligned with global trends.

D. Family Challenges

Family Challenges provides a comprehensive exploration of the struggles, values, and shared experiences that define family life (see Fig. 11). It captures the ways families navigate their everyday realities, from financial hardships to emotional resilience, and highlights the role of community, health, culture, and entertainment in shaping familial bonds and societal connections.

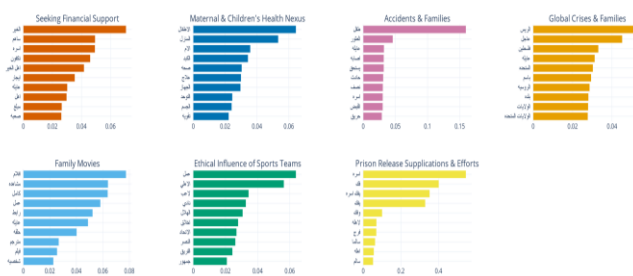


Fig. 11. c-TF-IDF Keyword scores (Family challenges).

At its foundation lies the Seeking Financial Support parameter, which reflects the financial challenges that many families face and their reliance on community assistance. As captured in one tweet, “A mother supporting her children asks for help paying an electricity bill, even if it's a small amount, minor contributions can make a significant difference...”.

Another tweet describes the severity of these struggles as “A dignified family with young children is in urgent need of essential supplies”, and “Requesting financial help for marriage; my father is deceased, and I am responsible for an entire family”. These narratives highlight the emotional and financial burdens shouldered by individuals and families, emphasizing the vital role of community empathy and charity.

Health and well-being are equally critical, as seen in Maternal & Children's Health Nexus. It underscores the importance of health for mothers and children as a cornerstone of family life. A tweet about prenatal care states, “A pregnant woman needs healthy sleep more than ever to support her own and her baby's health”, highlighting the value of self-care during pregnancy. The benefits of breastfeeding are echoed in another: “Breast milk is essential for the child's health”, pointing to the natural ways families support well-being. The role of nutrition is evident as well: “Zinc deficiency causes depression and attention issues in children”, reflecting the holistic measures required to ensure the health and happiness of families.

Accidents and Families brings to light the devastating impact of unforeseen events. One tweet recounts, “Teachers and three of their children were injured in a traffic accident”, demonstrating how sudden tragedies disrupt lives. The emotional toll of crimes targeting families is reflected in, “The punishment awaited for those responsible for kidnapping children and demanding ransom”, while the long-term pain of loss is encapsulated in, “My husband passed away nine years ago in a traffic accident”. These stories illustrate the fragility of family life and the enduring scars left by such tragedies.

Families are also deeply affected by broader external events, as captured in Global Crises and Families. A tweet highlights the lingering health impacts of the pandemic: “Children who have recovered from COVID-19 face an increased risk of developing type 1 or type 2 diabetes”, showing how crises can have lasting consequences. The global response to conflicts is reflected in, “The White House spokesperson mentioned a team for humanitarian aid to support the Ukrainian people”, showcasing solidarity and compassion. Meanwhile, the emotional toll of war is poignantly expressed: “O Lord, protect us from the evil of war, for its fuel is humans children, and women who bear no blame”, emphasizing the yearning for peace and the protection of the innocent.

The emotional resilience of families is evident in Prison Release Supplications and Efforts, where spiritual and community appeals play a vital role. One tweet conveys hope: “Grant freedom to a prisoner and reunite them with their family”, while another emphasizes communal responsibility: “Help ease the hardship of a prisoner your support is needed”. The profound emotional impact on loved ones is captured in, “If you can't feel the prisoner's pain, think of their mothers whose hearts can't bear this suffering”, showcasing the interconnectedness of familial bonds even during separation.

Cultural and moral influences are central to Ethical Influence of Sports Teams, where sports serve as a vehicle for teaching values and fostering unity. As one tweet observes, “Sports teach respect and morals” and “A true champion is defined by their conduct in defeat as much as in victory” highlighting the lessons learned from sportsmanship. Another tweet praises athletes as

role models: “This player is a great example; his ethics precede him on the field; he is a role model for behavior and skill” showing how sports inspire families to adopt principles of integrity and perseverance. However, the ethical influence of sports is not always positive. Some athletes and teams have been involved in controversies, with behaviors contradicting the values sports aim to promote. Issues such as unsportsmanlike conduct, corruption, or idolizing flawed role models can lead to negative influences, especially on young fans. Additionally, excessive sports engagement can overshadow family time or encourage an unhealthy competitive mindset. This duality emphasizes the need for families to engage critically with sports culture embracing their moral lessons while being mindful of its potential pitfalls.

Finally, Family Movies serve as a popular source of entertainment, but they can also have negative effects on children and families. Excessive screen time may reduce meaningful family interactions, and certain movie themes can introduce unrealistic expectations, harmful messages, or inappropriate behaviors. In some cases, these influences may lead to misunderstandings or conflicts within households. The potential for such negative impacts highlights the complex role family movies play in shaping family dynamics and children’s development.

To support families navigating societal challenges, policies should strengthen financial aid systems, improve maternal and child healthcare access, and address accident prevention through stricter safety regulations and post-incident support. Initiatives should mitigate the impacts of global crises by fostering resilience-building programs while promoting family-oriented cultural activities and sports to strengthen bonds. Advocacy for ethical influence through sports and reforms around prison release and reintegration are vital for societal cohesion. Future work should prioritize cross-disciplinary research, technology-driven solutions, cultural distinction, stakeholder collaboration, and continuous monitoring of policy outcomes to ensure adaptable, inclusive, and impactful interventions.

VI. DISCUSSION

This section presents a discussion of the overall findings of our research, along with policy recommendations and suggestions for future research.

The dynamics of family life in Saudi Arabia are profoundly shaped by the interplay between nurturing familial bonds, educational and career advancements, and the challenges families face. The Nurturing Families macro-parameter emphasizes the importance of strong family values, traditions, and parental roles in fostering emotional and social well-being. Simultaneously, Education and Career Development highlights how educational achievements and professional growth contribute to economic stability and overall family prosperity. Family Challenges sheds light on the various economic, health, and societal hurdles that families must navigate. Together, these macro-parameters create a comprehensive framework that illustrates how emotional support, economic empowerment, and resilience against challenges collectively enhance the well-being and progress of families in the context of modern societal changes and technological advancements.

A key synergy exists between family nurturing and educational and career advancement. Effective parental involvement, as highlighted in the Nurturing Families tweets, plays a critical role in fostering social and academic success. Parents who actively engage in their children’s education by providing guidance, maintaining discipline, and supporting educational pursuits create an environment conducive to learning and personal development. This support not only enhances students’ academic performance but also instills values essential for career readiness. In turn, success in education and career development contributes to economic stability, which strengthens family cohesion and overall well-being. Families with members who achieve higher educational and career milestones can invest more in their households, ensuring better living conditions, healthcare, and opportunities for future generations. This economic upliftment reduces financial stress, allowing families to focus more on nurturing relationships and maintaining strong familial bonds.

Cultural and spiritual practices serve as vital buffers that help families navigate economic, health, and societal challenges. Spiritual resilience fostered through practices such as healing prayers and collective religious activities, provides emotional and psychological support that enables families to cope with stressors related to education and career pressures. This resilience fosters a stable environment where individuals can pursue their educational and professional aspirations without being overwhelmed by external challenges. Additionally, cultural traditions such as Eid Gatherings and Condolence Gatherings strengthen family unity, ensuring that during times of financial hardship or global crises, families remain cohesive and supportive. This unity is crucial for maintaining a stable home environment, which is essential for both educational success and career advancement.

Economic stability is a cornerstone that underpins both overcoming challenges and enhancing family life. Quality education provides individuals with the skills and knowledge required to obtain better employment opportunities, leading to economic empowerment. This empowerment not only improves individual livelihoods but also enhances the collective well-being of families, enabling them to invest in health, education, and other essential areas. Conversely, economic hardships such as struggling to meet basic needs or seeking financial support directly impact family dynamics. Families facing financial stress may experience increased tension and reduced cohesion, undermining the nurturing environment necessary for educational and career success. Addressing these financial challenges through supportive policies and community initiatives is essential for maintaining family stability and enabling continued progress.

Smart home technologies emerge as pivotal tools that bridge family support and modern challenges. These technologies can enhance family cohesion and communication by facilitating better interactions among family members, ensuring that despite busy schedules related to education and career, families remain connected. Smart devices for managing daily schedules, facilitating virtual gatherings, and automating household tasks free up time for meaningful interactions, thereby strengthening family bonds. Additionally, digital platforms integrated into smart homes provide access to educational resources, online

learning tools, and career planning applications, supporting continuous learning and professional growth even in the face of external challenges such as global crises or economic downturns. Furthermore, smart home technologies can enhance health monitoring and safety within the household, addressing some of the family challenges related to maternal and child health or accidents by providing features such as health tracking devices, emergency alerts, and remote health consultations.

Societal influences, including media representations and ethical considerations from sports, intersect with family values and education, shaping the moral and ethical framework within which families operate. Positive representations in media and sports can inspire individuals and families to pursue higher educational and career goals, while negative portrayals may create challenges that families must navigate collectively. This underscores the need for strong family support systems to uphold desired values and aspirations. Additionally, societal expectations and cultural norms significantly influence career choices and educational pursuits, with families often acting as mediators to help individuals align their aspirations with societal expectations, thereby shaping their educational and professional trajectories.

The interconnected nature of economic instability, health concerns, external crises, and cultural influences necessitates a holistic approach to supporting families. Integrated policy reforms that simultaneously address educational improvements, economic support, and healthcare enhancements can create a more supportive environment for families. Community initiatives, supported by technological advancements, can provide comprehensive support through programs offering financial counseling, healthcare services, and educational support, combined with smart home technologies to create resilient family structures capable of withstanding various challenges.

Several key insights emerge from the cross-parameter analysis. Firstly, a Virtuous Cycle of Support and Progress is evident, where strong family nurturing leads to educational and career success, which in turn enhances economic stability and further strengthens family bonds, creating a positive feedback loop that fosters continuous growth and resilience. Additionally, Cultural and Spiritual Resilience plays a crucial role, as the integration of cultural traditions and spiritual practices provides emotional and psychological resilience, enabling families to effectively navigate both personal and societal challenges while maintaining focus on educational and career goals. Economic Empowerment Facilitates Comprehensive Well-Being by achieving economic stability through education and career advancement, which not only improves individual livelihoods but also enhances the overall well-being of families, reducing financial stress and enabling investment in family health and cohesion. Furthermore, Technological Integration as a Support Mechanism is significant, with smart home technologies acting as enablers that support both family nurturing and educational/career development by facilitating better communication, access to resources, and safety measures, thereby bridging the gap between traditional family values and modern challenges. The analysis also highlights Policy and Community Synergy, where effective policy measures addressing multiple facets of family life, such as education,

economic support, and healthcare, combined with community initiatives, create a robust support system that enhances family resilience and progress. Lastly, Adaptive Capacity Through Integrated Support underscores that families benefiting from integrated support systems, combining strong familial bonds, educational opportunities, economic stability, and cultural resilience, demonstrate greater adaptive capacity in the face of challenges, ensuring sustained well-being and development.

A. Policy Recommendations and Future Work

Future research and implementation should focus on developing integrated support programs that simultaneously address family cohesion, educational support, and economic empowerment to maximize the synergistic benefits identified in the analysis. Additionally, it is crucial to leverage smart home technologies for family support by investing in systems that facilitate both emotional well-being and practical needs, such as virtual family gatherings, remote education tools, and health monitoring and management systems. Promoting cultural and spiritual engagement should be encouraged to strengthen family bonds and provide resilience against external challenges. A key policy recommendation is to formulate and implement policies centered on holistic family welfare, recognizing and supporting the interconnected nature of family well-being, education, and economic stability. These policies should ensure that interventions are comprehensive and multifaceted, addressing various aspects of family life simultaneously. Furthermore, conducting comparative cross-regional studies is recommended to identify unique cultural factors and best practices that can be adapted to the Saudi context for enhancing family welfare. Enhancing technological literacy and accessibility is also vital to ensure that families have access to and are proficient in using smart home technologies, thereby maximizing their potential to support family cohesion, education, and career development. Additionally, policies should incentivize the adoption of smart technologies in households through subsidies or training programs to bridge the digital divide. By embracing a comprehensive and integrated approach, stakeholders can better support families in Saudi Arabia, leveraging the strengths of cultural traditions and modern advancements to foster environments where families can thrive both emotionally and economically.

VII. CONCLUSION

This study provides a comprehensive analysis of Smart Homes & Families by examining 1,174,912 Arabic tweets from Saudi Arabia to uncover societal perceptions, challenges, and expectations. The findings highlight the interconnected roles of nurturing familial bonds, educational and career development, and overcoming family challenges in shaping family well-being. The analysis revealed how parental involvement, economic stability, and cultural and spiritual practices contribute to educational success and familial cohesion. Furthermore, smart home technologies emerged as key enablers, supporting family communication, education, healthcare, and overall resilience. The study also emphasized the influence of media representations, societal expectations, and cultural norms on family dynamics.

Through a comparative analysis of AraBERT, MarBERT, and DistilBERT, the research identified AraBERT as the most

effective model for analyzing Arabic Twitter data. The findings are supported by detailed thematic analyses, a structured taxonomy, and policy recommendations aimed at enhancing family well-being and societal sustainability. This work advances previous research [26] by offering deeper thematic insights, improved data analysis methodologies, and practical recommendations, paving the way for future studies on the role of smart technologies in supporting family resilience and societal development in Saudi Arabia.

ACKNOWLEDGMENT

“This article is derived from a research grant funded by the Research, Development, and Innovation Authority (RDIA), Kingdom of Saudi Arabia, with grant number 12615-iu-2023-IU-R-2-1-EI-.”

REFERENCES

- [1] T. Yigitcanlar et al., “Artificial Intelligence Technologies and Related Urban Planning and Development Concepts: How Are They Perceived and Utilized in Australia?,” *J. Open Innov. Technol. Mark. Complex.*, vol. 6, no. 4, p. 187, Dec. 2020, doi: 10.3390/joitmc6040187.
- [2] R. Mehmood, A. Sheikh, C. Catlett, and I. Chlamtac, “Editorial: Smart Societies, Infrastructure, Systems, Technologies, and Applications,” *Mobile Networks and Applications*, vol. 28, no. 2, Springer, pp. 598–602, May 03, 2023, doi: 10.1007/s11036-022-01990-y.
- [3] S. Prasad, G. Hossain, A. Goyal, A. Bhan, and S. Bhattacharya, “Smart home health monitoring system for predicting type 2 diabetes and hypertension,” *J. King Saud Univ. - Comput. Inf. Sci.*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2020.01.010.
- [4] Q. Lu, Z. Zhang, and S. Lü, “Home energy management in smart households: Optimal appliance scheduling model with photovoltaic energy storage system,” *Energy Reports*, vol. 6, pp. 2450–2462, Nov. 2020, doi: 10.1016/J.EGYR.2020.09.001.
- [5] G. Alexakis, S. Panagiotakis, A. Fragakakis, E. Markakis, and K. Vassilakis, “Control of smart home operations using natural language processing, voice recognition and iot technologies in a multi-tier architecture,” *Designs*, vol. 3, no. 3, pp. 1–18, 2019, doi: 10.3390/designs3030032.
- [6] W. Li, T. Yigitcanlar, A. Liu, and I. Erol, “Mapping two decades of smart home research: A systematic scientometric analysis,” *Technol. Forecast. Soc. Change*, vol. 179, p. 121676, Jun. 2022, doi: 10.1016/J.TECHFORE.2022.121676.
- [7] N. Janbi, R. Mehmood, I. Katib, A. Albeshri, J. M. Corchado, and T. Yigitcanlar, “Imtidad: A Reference Architecture and a Case Study on Developing Distributed AI Services for Skin Disease Diagnosis over Cloud, Fog and Edge,” *Sensors*, vol. 22, no. 5, p. 1854, Feb. 2022, doi: 10.3390/s22051854.
- [8] N. Janbi, I. Katib, A. Albeshri, and R. Mehmood, “Distributed Artificial Intelligence-as-a-Service (DAIaaS) for Smarter IoE and 6G Environments,” *Sensors*, vol. 20, no. 20, p. 5796, Oct. 2020, doi: 10.3390/s20205796.
- [9] S. Sehgal, H. Sharma, and A. Anand, “Smart and Context-Aware System employing Emotions Recognition,” pp. 1–8, 2021, doi: 10.1109/incet51464.2021.9456356.
- [10] M. Shuai, N. Yu, H. Wang, and L. Xiong, “Anonymous authentication scheme for smart home environment with provable security,” *Comput. Secur.*, vol. 86, pp. 132–146, Sep. 2019, doi: 10.1016/J.COSE.2019.06.002.
- [11] P. J. Rani, J. Bakthakumar, B. P. Kumar, U. P. Kumar, and S. Kumar, “Voice controlled home automation system using natural language processing (NLP) and internet of things (IoT),” *ICONSTEM 2017 - Proc. 3rd IEEE Int. Conf. Sci. Technol. Eng. Manag.*, vol. 2018-Janua, pp. 368–373, 2017, doi: 10.1109/ICONSTEM.2017.8261311.
- [12] W. Choi, J. Kim, S. E. Lee, and E. Park, “Smart home and internet of things: A bibliometric study,” *J. Clean. Prod.*, vol. 301, p. 126908, Jun. 2021, doi: 10.1016/j.jclepro.2021.126908.
- [13] D. Marikyan, S. Papagiannidis, and E. Alamanos, “A systematic review of the smart home literature: A user perspective,” *Technol. Forecast. Soc. Change*, vol. 138, pp. 139–154, Jan. 2019, doi: 10.1016/j.techfore.2018.08.015.
- [14] J. F. Defranco and M. Kassab, “Smart Home Research Themes: An Analysis and Taxonomy,” in *Procedia Computer Science*, Jan. 2021, vol. 185, pp. 91–100, doi: 10.1016/j.procs.2021.05.010.
- [15] S. Pira, “The social issues of smart home: a review of four European cities’ experiences,” *Eur. J. Futur. Res.*, vol. 9, no. 1, 2021, doi: 10.1186/s40309-021-00173-4.
- [16] A. Singh, J. Kumar, A. Jha, and S. Purbey, “Bibliometric Analysis of Home Health and Internet of Health Things (IoHT),” *Lect. Notes Electr. Eng.*, vol. 776, pp. 75–88, 2022, doi: 10.1007/978-981-16-2911-2_9/COVER/.
- [17] P. Li, Y. Lu, D. Yan, J. Xiao, and H. Wu, “Scientometric mapping of smart building research: Towards a framework of human-cyber-physical system (HCPS),” *Autom. Constr.*, vol. 129, p. 103776, Sep. 2021, doi: 10.1016/J.AUTCON.2021.103776.
- [18] M. Abdul-Mageed, A. R. Elmadany, and E. M. B. Nagoudi, “ARBERT & MARBERT: Deep bidirectional transformers for Arabic,” *ACL-IJCNLP 2021 - 59th Annu. Meet. Assoc. Comput. Linguist. 11th Int. Jt. Conf. Nat. Lang. Process. Proc. Conf.*, no. ii, pp. 7088–7105, 2021, doi: 10.18653/v1/2021.acl-long.551.
- [19] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” Oct. 2019.
- [20] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” 2018.
- [21] L. McInnes, J. Healy, and S. Astels, “hdbSCAN: Hierarchical density based clustering,” *J. Open Source Softw.*, vol. 2, no. 11, p. 205, Mar. 2017, doi: 10.21105/joss.00205.
- [22] M. Grootendorst, “GitHub - MaartenGr/cTFIDF: Creating class-based TF-IDF matrices.”
- [23] “Histograms — Matplotlib 3.5.2 documentation.”
- [24] “seaborn.heatmap — seaborn 0.11.2 documentation.”
- [25] “Plotly: Low-Code Data App Development.”
- [26] E. Alqahtani, N. Janbi, S. Sharaf, and R. Mehmood, “Smart Homes and Families to Enable Sustainable Societies: A Data-Driven Approach for Multi-Perspective Parameter Discovery Using BERT Modelling,” *Sustainability*, vol. 14, no. 20, p. 13534, Oct. 2022, doi: 10.3390/SU142013534.
- [27] K. Gram-Hanssen and S. J. Darby, “‘Home is where the smart is’? Evaluating smart home research and approaches against the concept of home,” *Energy Res. Soc. Sci.*, vol. 37, pp. 94–101, Mar. 2018, doi: 10.1016/J.ERSS.2017.09.037.
- [28] C. Després, “The Meaning of Home: Literature Review and Directions for Future Research and Theoretical Development,” *J. Archit. Plann. Res.*, vol. 8, no. 2, pp. 96–115, 1991, Accessed: Aug. 01, 2022. [Online]. Available: <https://www.jstor.org/stable/43029026>.
- [29] E. Mitty and S. Flores, “There’s No Place Like Home,” *Geriatr. Nurs. (Minneapolis)*, vol. 30, no. 2, pp. 126–129, 2009, doi: <https://doi.org/sdl.idm.oclc.org/10.1016/j.gerinurse.2009.01.004>.
- [30] D. Hatcher, E. Chang, V. Schmied, and S. Garrido, “Exploring the Perspectives of Older People on the Concept of Home,” *J. Aging Res.*, vol. 2019, 2019, doi: 10.1155/2019/2679680.
- [31] F. A. Lewin, “The Meaning of Home among Elderly Immigrants: Directions for Future Research and Theoretical Development,” <http://dx.doi.org/10.1080/02673030120049715>, vol. 16, no. 3, pp. 353–370, 2010, doi: 10.1080/02673030120049715.
- [32] S. Alotaibi, R. Mehmood, I. Katib, O. Rana, and A. Albeshri, “Schaa: A Big Data Analytics Tool for Healthcare Symptoms and Diseases Detection Using Twitter, Apache Spark, and Machine Learning,” *Appl. Sci.*, vol. 10, no. 4, p. 1398, Feb. 2020, doi: 10.3390/app10041398.
- [33] E. Alomari, I. Katib, A. Albeshri, T. Yigitcanlar, and R. Mehmood, “Iktishaf+: A Big Data Tool with Automatic Labeling for Road Traffic Social Sensing and Event Detection Using Distributed Machine Learning,” *Sensors*, vol. 21, no. 9, p. 2993, Apr. 2021, doi: 10.3390/s21092993.

- [34] J. R. Saura, D. Palacios-Marqués, and D. Ribeiro-Soriano, "Using data mining techniques to explore security issues in smart living environments in Twitter," *Comput. Commun.*, vol. 179, pp. 285–295, Nov. 2021, doi: 10.1016/J.COMCOM.2021.08.021.
- [35] Y. Su, A. Venkat, Y. Yadav, L. B. Puglisi, and S. J. Fodeh, "Twitter-based analysis reveals differential COVID-19 concerns across areas with socioeconomic disparities," *Comput. Biol. Med.*, vol. 132, p. 104336, May 2021, doi: 10.1016/J.COMPBIOMED.2021.104336.
- [36] M. Abdulaziz, A. Alotaibi, M. Alsolamy, and A. Alabbas, "Topic based Sentiment Analysis for COVID-19 Tweets," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 1, pp. 626–636, 2021, doi: 10.14569/IJACSA.2021.0120172.
- [37] S. Alswedani, R. Mehmood, and I. Katib, "Sustainable Participatory Governance: Data-Driven Discovery of Parameters for Planning Online and In-Class Education in Saudi Arabia During COVID-19," *Front. Sustain. Cities*, vol. 4, p. 97, Jul. 2022, doi: 10.3389/FRSC.2022.871171/BIBTEX.
- [38] S. Alswedani, R. Mehmood, I. Katib, and S. M. Altowaijri, "Psychological Health and Drugs: Data-Driven Discovery of Causes, Treatments, Effects, and Abuses," *Toxics* 2023, Vol. 11, Page 287, vol. 11, no. 3, p. 287, Mar. 2023, doi: 10.3390/TOXICS11030287.
- [39] A. S. Alammery, "BERT Models for Arabic Text Classification: A Systematic Review," *Appl. Sci.*, vol. 12, no. 11, p. 20, 2022, doi: 10.3390/app12115720.
- [40] M. Röder, A. Both, and A. Hinneburg, "Exploring the space of topic coherence measures," *WSDM 2015 - Proc. 8th ACM Int. Conf. Web Search Data Min.*, pp. 399–408, 2015, doi: 10.1145/2684822.2685324.
- [41] "Models.coherencemodel – Topic coherence pipeline — gensim."