Advances in Artificial Intelligence and Machine Learning 3 (3) 1460-1481 Received 15-05-2023; Accepted 23-09-2023; Published 30-09-2023

Deploying Change Modeling to Study the Evolution of COVID-19 Related Menstrual Health Issues

Sarabjeet Kaur Kochhar skaur@ip.du.ac.in
Department of Computer Science
Indraprastha College for Women
University of Delhi
New Delhi, India

Anamika Gupta anamikargupta@sscbsdu.ac.in
Department of Computer Science
Shaheed Sukhdev College of Business Studies
University of Delhi
New Delhi, India

Corresponding Author: Anamika Gupta

Abstract

The menstrual health of women serves as a critical marker of their overall health and quality of life. Considering the established link between irregular menstrual cycles, associated issues, and cardiovascular diseases, along with the potential implications for premature mortality, there is an urgent need to prioritize research efforts dedicated to assessing the impact of the COVID-19 pandemic on women's menstrual health. It's surprising to observe that the machine-learning community has largely overlooked this issue, which significantly impacts nearly half of the global population. The work presented in this paper utilizes change modeling to investigate the evolution of COVID-19-related menstrual health issues over time, as gleaned from tweets. Change modeling involves monitoring shifts in the acquired knowledge over time and creating innovative change measures or metrics to organize and potentially restructure this information. This approach to tracking changes offers a fresh, consolidated, and abstract perspective for examining the collected data, opening up new methods for analysis and inference. An additional valuable aspect of change modeling is its capacity to enhance the understanding of results, thereby increasing their practical utility. The outcomes of the framework presented here consist of innovative, straightforward, and easily understandable metrics capable of capturing changes from multiple perspectives. The primary benefit lies in the fact that these simple yet elegant change metrics simplify the decision-making process, providing a rapid and concise, one-glance overview of the metrics. The framework also serves as a foundation for generating only a limited selection of well-defined association links derived from discussions about menstruation on Twitter. This activity aligns with the goal of producing well-qualified, comprehensible, and contextually relevant knowledge, and provides advantages in terms of computational efficiency and storage space.

Citation: Sarabjeet kaur kochhar and Anamika Gupta. Deploying Change Modeling to Study the Evolution of COVID-19 Related Menstrual Health Issues. Advances in Artificial Intelligence and Machine Learning. 2023;3(3):85.

Keywords: COVID-19, Menstrual health, Change modeling, Abstract, Association rule generation.

1. INTRODUCTION

Social media platforms serve as forums where people from different strata of society across the world voice their unhindered opinions on a variety of subjects. Such platforms have been utilized to mine patterns that reflect the viewpoints, thought processes, and attitudes of people. Sometimes the knowledge obtained from social media messages highlights the emergent, actionable needs of the people. The rising anxiety and several problems, showcased by the tweets related to menstruation during the pandemic, serve as a ready reckoner for the same.

It is known that people worldwide were exposed to grave and persistent psychosocial stressors due to the COVID-19 pandemic. Psychological studies have established that women have been more susceptible to the stress caused by the COVID-19 pandemic than men, thereby bringing about changes in their menstrual cycles akin to those caused by other, extremely stressful situations like war, displacement, natural disasters, and famine [1]. Stress due to panic, alarm, terror, and distress affects the hypothalamic-pituitary-gonadal (HPG) and hypothalamic-adrenal axis, leading to menstrual cycle irregularities, dysmenorrhea, pre-menstrual symptoms, and menorrhagia [1–3]. A study recently reported that the elevated levels of anxiety and stress triggered by the COVID-19 outbreak were substantial enough to impact the attributes of the menstrual cycle in women [4].

The menstrual health of a woman is an important indicator of her overall well-being and quality of life. Menstrual problems add a lot of economic burden to women, their families, and society in general. Irregular cycles and menstrual problems have even been linked to cardiovascular diseases and premature mortality [5, 6]. Research to study the effect of the COVID-19 pandemic on women's menstrual health is therefore an imperative public health initiative and the need of the hour. It is disappointing that only a handful of medical studies have been conducted so far in this important area. In fact, the few survey-based studies carried out by the medical fraternity have also pointed out the need for more serious work to be done in this direction [6–9]. It is noteworthy that there has been almost no work done by the machine learning community in the aforementioned area of women's health. New age technologies like data mining, artificial intelligence, and machine learning have the prowess to algorithmically harness gigantic datasets, sophisticated pattern detection, and complement the medical fraternity by presenting them with the consolidated, understandable knowledge derived from data of millions of people, in the form of patterns, trends, etc. This ready-to-be-leveraged knowledge can be used by domain experts to draw conclusions and inferences.

A breakthrough work in this direction by Kochhar et al. has attempted to break the taboo and employs a blend of supervised and unsupervised data mining techniques to study menstrual health [10]. The knowledge differentiation-based framework facilitates the discovery and analysis of knowledge from multiple perspectives, enabling its consolidation, comprehension, and actionability.

The work presented in this paper carries forward the aforementioned work by presenting a framework to carry out the following tasks:

1. Deploys change modeling to study, over time, the COVID-19-related menstrual health issues from tweets. Change modeling is the process of monitoring changes in the mined knowledge, over time, and designing novel measures or metrics to consolidate and possibly restructure the generated knowledge. This modeling of changes provides a new, consolidated, abstract view of looking at the mined information and allows novel ways of analyzing it and drawing inferences. Another important benefit of change modeling is that it promotes comprehension of the results, thereby improving their actionability. The deployment of change modeling is detailed in Section 4.

2. Forms basis for the generation of a small number of selected, well-labeled association linkages from the menstrual talk on Twitter. This falls under the ambit of generating qualified, understandable, context-aware knowledge. Gauging the interestingness of discovered knowledge by its understandability and actionability has been studied and stressed by many data mining research endeavors [11–13]. The generation of a very large number of association rules means a voluminous output that requires considerable time and effort to be sifted. Reduction in the number of association rules is thus highly desirable. If the generated rules are also labeled, they provide a rich context that helps in their comprehension and deployment. Further details of this task are presented in Section 5.3.

   The outcomes of the presented framework encompass novel, simple and intuitive, metrics that capture changes from multiple perspectives. These thoughtfully designed change metrics facilitate decision-making in a single glance. The framework also set a platform for generation of a reduced number of relevant, well-categorized association rules, offering the benefits of computation time and space.

The paper is structured in the following manner: Section 2 provides a review of related work. Section 3 introduces the general framework proposed in this paper. Section 4 outlines the change modeling framework. Section 5 delves into the details of implementation. Section 6 presents the results of the implemented framework accompanied by a discussion of noteworthy findings. Section 7 offers conclusions, summarizes the paper and lays down avenues for future work.

2. RELATED WORK

This section provides an overview of related research that closely aligns with our study. To the best of our knowledge, this paper represents one of the handful machine learning works aimed at addressing the previously unexplored topic of studying changes in women's menstrual health in relationship with the COVID-19 pandemic.

2.1 Menstrual Health and COVID

The machine learning community has not contributed much to this important facet that affects almost half the population of the world. Resultantly, largely studies conducted by the medical community form part of this section.

Nazir et al. present a questionnaire-based cross-sectional design review of the latest findings regarding the factors influencing the occurrence of menstrual issues following COVID-19 vaccination, encompassing a total of 78,138 patients [14]. The study concludes that a substantial proportion of women (52.05%) have reported experiencing menstrual abnormalities such as heavy menstrual bleeding, infrequent periods, and painful periods after receiving the COVID-19 vaccine. However, most of the cross-sectional studies included in the literature review were unable to establish a direct causal link between menstrual irregularities and the status of COVID-19 vaccination.

Another literature review including three studies utilizing the Medline, Scopus, and Cochrane Library databases was carried by Lebar et al. [8]. The review concludes that SARS-CoV-2 infection can result in changes in menstrual volume and alterations in the length of menstrual cycles. Notably, an elongated menstrual cycle was the most frequently reported menstrual irregularity in these studies, and women predominantly reported reduced menstrual volume. A similar survey conducted by chao et al. re-enforce these findings [15]. Demir et al. uncover a noteworthy correlation between the heightened levels of anxiety and stress stemming from the COVID-19 crisis and alterations in the menstrual cycle patterns among the surveyed women [4].

Muhaidat et al. conducted a cross-sectional survey via social media between July 2021 and August 2021 that concludes that females who have received COVID-19 vaccines may encounter prolonged menstruation or extended menstrual cycle durations, among other potential abnormalities, though these symptoms might be transient and self-limited [9]. A cross-sectional online investigation was carried out as a component of a project called “Effect of Vaccination against SARS-CoV-2 on the Menstrual Cycle (EVA Project)” [16]. Of the surveyed women, 78% reported experiencing premenstrual changes including high fatigue, abdominal bloating, emotional irritability, feelings of sadness, headaches, and breast pain. Additionally, the most frequently reported menstrual changes encompassed heavier menstrual bleeding, increased pain, and delayed menstruation. A detailed review of some more works in the stated problem domain has been presented in [10].

It is pertinent to note here that almost all the studies reviewed above underscore the fact that menstrual health and its related alterations are not adequately addressed in fundamental and applied research efforts.

2.2 Change Modeling

To the best of our knowledge no work has been devoted to change modeling except the works reported in [17], and [18], which have formed the basis of extension for the work presented in this paper. Most of the related works in the machine learning have limited themselves to change detection or evolution studies, areas which have received considerable attention since the last three decades, and are reviewed in this section.

Change detection algorithms are typically designed to identify alterations that occur at a specific point in time, making them effective for static databases. However, in the context of continuously incoming data, it becomes imperative to monitor the progression and evolution of these changes over time. This is the domain of evolution studies. Nevertheless, the existing body of research in evolution studies often lacks a focus on designing systems that produce results that are easy to comprehend, deploy, and potentially analyzable across different levels of abstraction. This is where change modeling steps in to bridge this gap.

Haque et al. use the concept of sliding windows to detect concept drift in classification models constructed from data streams and update them [19]. Gupta et al. proposed an incremental approach for discovery of frequent itemsets from the data streams [20, 21], and the generation of non-redundant association rules using Formal Concept Analysis [22], a branch of mathematics based on concepts and its hierarchy. Further, multi-objective associative classifiers were discovered using the genetic algorithms [23]. A spatial-temporal data mining framework for change detection using APRIORI-based frequent itemset mining algorithm has been deployed to comprehend alterations in the attributes of human-caused wildfires in Western Colorado between 1992 and 2015, taking into account both their geographic distribution and spatial surroundings [24]. An approach for mining frequent patterns in evolving data streams, referred to as VSWCDD (Variable Sliding Window-Concept Drift Detection), that adapts to changing data patterns by employing a variable sliding window, dynamically adjusting its size based on the occurrence of concept drift in the data stream has been presented by Yin et al. [25]. A study that uses Support Vector Machines (SVM) and Genetic Algorithms (GA) to enhance change for detecting changes in satellite images is presented by Pati et al. [24]. Satellite photos from 1 January 2019 to 4 March 2020 were utilized to conduct a supervised classification of LULC (Land Use/Land Cover) using SVM classifier to detect changes in the land cover mapping in Puri District, Odisha, India by Swetanisha et al. [26].

3. RESEARCH METHODOLOGY

In this section, we first lay down the details of the change modeling framework proposed in this paper. The presented change model is an extension of the knowledge differentiation framework proposed by Kochhar et al. [10], which is summarized next. The change modeling framework proposed in this paper is presented in FIGURE 1. It deploys a focusing unit that captures user subjectivity and allows the user to provide parameters that place the control of monitoring and modeling in the user's hands. The dimensions along which user subjectivity is captured are outlined in Section 4. The change detection unit receives the support of itemsets to be monitored and modeled from the existing knowledge differentiation framework presented in FIGURE 2. It is responsible for finding out the changes in support of these itemsets over time. The consolidation and restructuring unit observes and aggregates the changes for predefined time granularities via consolidation windows and restructures them by applying appropriate consolidation functions. The unit generates novel change metrics at different levels of abstraction. Each level of abstraction provides a new way of looking at the changes and aids in advanced analytics and inferencing. The user can look at the modeled changes and their corresponding labels that indicate their categories. It is the user's decision to generate association rules only for the categories of itemsets that interest him. Hence, a limited number of well-labeled and categorized rules are generated, that are of definite interest to the user.

The knowledge differentiation framework scrapes and preprocesses the COVID-19 tweets talking about women's' menstrual health. The framework then separates the facts from the myths being circulated on Twitter about women's menstrual problems in order to understand public's authentic opinions and problems and the myths being populated about the same. Association rules are directly generated from the selected categories of tweets. The framework presented in this paper extends the

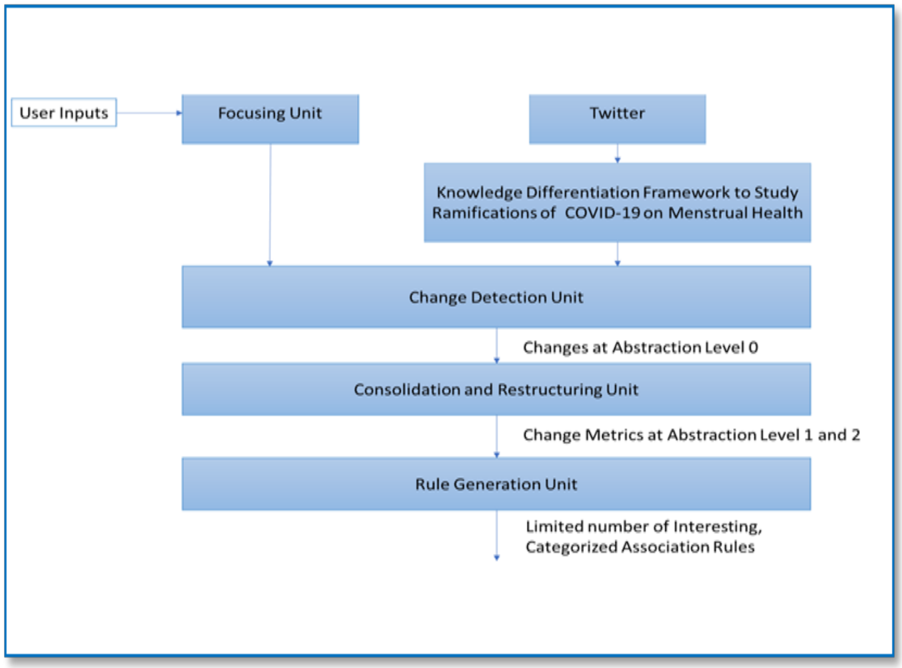

Figure 1: Change Modeling Framework to Study Evolution of Menstrual Changes in Relation to COVID-19

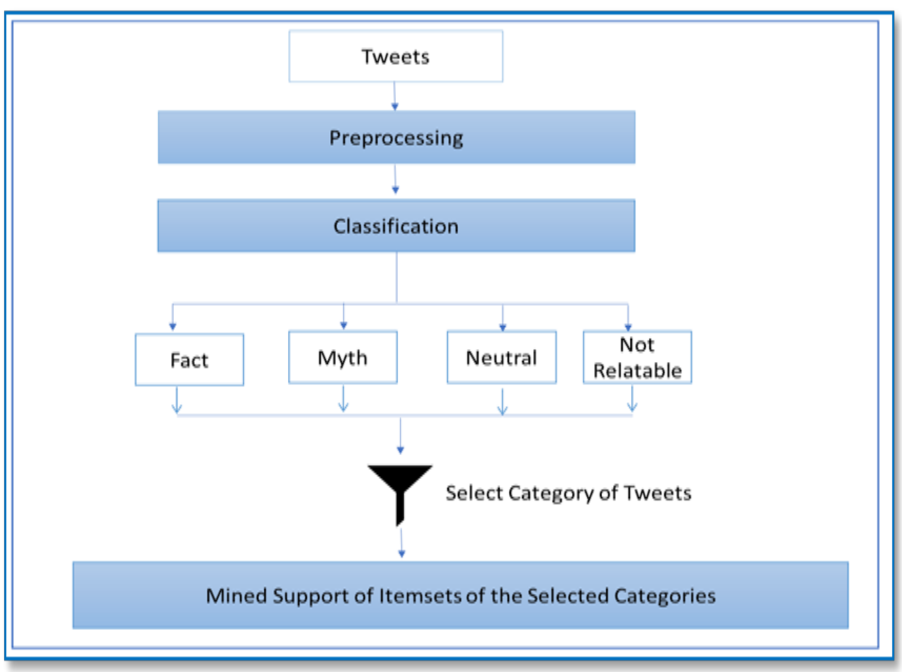

Figure 2: Knowledge Differentiation Framework to study COVID-19 Ramifications on Women's Menstrual Health

above framework by introducing monitoring and modeling. Novel change metrics are proposed that are computed over consolidation windows of different time granularities. These metrics provide a new way of analyzing and inferring the changing support behaviors. The association rules generated using the current framework are much smaller than the earlier framework and more comprehensible.

4. MODELING CHANGES

Monitoring changes is very important to study the evolution of trends. For example, in the context of menstrual talk on social media, a fluctuation in the support of words used indicates a likely change in the popularity of the words occurring in menses related talk. If this change in support is monitored over time, it may lead to the detection of a change in underlying pattern. Study of changing patterns is important to understand evolution. For instance, in the current context, a change in the topics people talk about may indicate a change in the menstrual problems being faced by people or a change in their perception of the same. In this section, the general concept of a change model is laid down first (Subsection 4.1), followed by the framework for modeling changes in menstrual problems related to COVID-19 reported by people on Twitter, proposed in this paper (Subsection 4.2).

4.1 Change Model

Restructuring of the changes, at multiple levels of abstraction, possibly by designing domain specific novel metrics over time is called change modeling. Expressing the changes detected over time in different ways, for example by using novel metrics helps in analyzing them in novel ways. A change model is essentially defined by four facets viz. user subjectivity, change detection, abstraction, and consolidation, as depicted by FIGURE 3.

A user, such as a domain expert, asserts substantial influence over the other three facets of change modeling. The change detection strategy, the perspectives with which they need to be viewed, and consolidation mechanisms to yield these views must be chosen in accordance with the user requirements. A change detection algorithm is required to systematically identify the changes in the mined knowledge over time. Such an algorithm may be selected or designed in accordance with the specific domain requirements, under the guidance of an experienced user. Abstraction allows us to look at the detected changes from multiple perspectives, providing novel, useful, deeper insights. Each level of abstraction involves consolidating the detected changes over a consolidation window and caters to a different semantic requirement. A consolidation window represents a time period over which the changes are combined, using a consolidation function. A consolidation function may be viewed as one or more metrics to restructure the changes and may be designed according to the domain. It may be applied over a consolidation window at a lower level of abstraction to generate lower order changes and over consolidation windows at higher level of abstraction to yield higher order changes. Moreover, different consolidation window sizes may be used at different levels of abstraction.

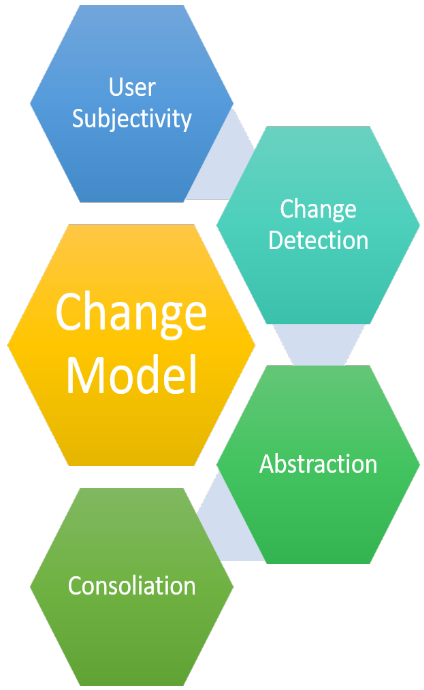

Figure 3: Facets of a Change Model

4.2 Framework for Modeling Changes in Menstrual Problems related to COVID-19

In this section, we present a change model to study changes in the support behaviour of menstrual itemsets related to COVID-19. The framework proposed in this paper is an application of the general change model presented by Bhatnagar et al. [17], to the framework to study COVID-19 ramifications on Women's Menstrual Health by Kochhar et al [10]. Formulation of the change model for modeling changes in menstrual problems related to COVID-19 with respect to its four facets is presented in the following subsections.

4.2.1 User subjectivity

The number of itemsets generated by an association rule mining algorithm is generally very large. Monitoring every itemset generated is neither practically achievable nor of significant interest from a computational perspective. In fact, focused monitoring is a very important requirement of most applications that work with itemsets that are mined in real time. Accordingly, for the work undertaken in this paper, we propose the following mechanism to focus monitoring by capturing user subjectivity along the two dimensions:

1. The Monitoring Space Ms [18], is specified by a user as a set of itemsets interesting for him to monitor. In this paper, we allow users to define the category of itemsets to be monitored and the particular itemsets in the specified category to be monitored.

2. Support Space [0,1] is the second dimension along which user subjectivity is captured. The user is allowed to specify multiple support thresholds which may be meaningful for him. Emphasis on a single support threshold is not meaningful for many applications [27]. In the present framework, we work with two support thresholds, which partition the support space into three partitions. The three partitions so formed represent the low support (Infrequent) itemsets, itemsets with medium support (Subfrequent itemsets) and high support (Frequent) itemsets can be seen in FIGURE 4. The itemsets selected for Ms in Step 1 are monitored along the partitioned support space.

4.2.2 Change detection

The changes in the support of itemsets are detected by capturing their movement across the three partitions of the support space, shown in FIGURE 4. This mechanism ensures discounting the very small changes, and considering the significant support changes that place an itemset into a different partition altogether. Let a Partition *i* be defined as an interval [S_l^i, S_h^i [, where S_l^i and S_h^i denote the lower and higher support boundaries of the partition respectively.

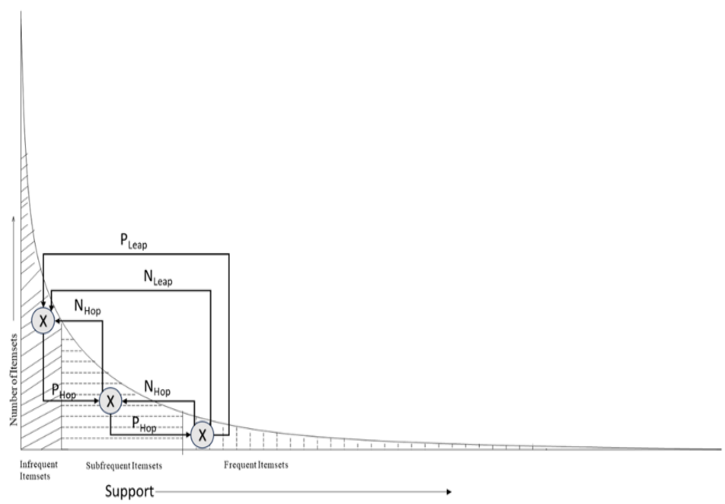

Figure 4: Hops and Leaps in the Partitioned Support Space

The Membership M_x^t of an itemset *x* at time instant *t* with support s_x^t can be defined as a mapping from the Set of Real numbers <=1 to the set of Positive Integers <=p, where p is the number of partitions,

$$R_1^+ \rightarrow Z_p^+$$

such that:

$$M_x^t = \{P_i | S_l^i <= s_x^t < S_h^i\}$$

Then the following metrics can be proposed to capture the support changes of itemsets:

Hops A Hop is defined as the transition of an itemset from a support partition to its neighboring support partition (FIGURE 4). Intuitively, a hop captures a significant but small transition in the support of an itemset. A hop is considered positive if the itemset moves from lower support partition to the neighboring higher partition and negative otherwise. If s_x^{t1} denotes the support of an itemset x at time $t1$ such that Membership of x at time $t1$, $M_x^{t1} = P_i$ and s_x^{t2} denotes the support of an itemset x at time $t2$ s.t. Membership of x at time $t2$, $M_x^{t2} = P_j$, where P_i and P_j denote two partitions of the support space, s.t. $t1 < t2$ and $i < j$, then a positive hop, P_{Hop} is defined as:

$$P_{\text{Hop}} = (j - i == 1)?1 : 0$$

and a negative hop, N_{Hop} is defined as:

$$N_{\text{Hop}} = (j - i == -1)?1 : 0$$

Leaps A Leap is defined as the transition of an itemset from a support partition to its non-neighboring support partition (FIGURE 4). Intuitively, a leap captures a significant and large transition in the support of an itemset. A leap is considered positive if the itemset moves from lower support partition to a non-neighboring higher partition and negative otherwise. If s_x^{t1} denotes the support of an itemset x at time $t1$ s.t. Membership of x at time $t1$, $M_x^{t1} = P_i$ and s_x^{t2} denotes the support of an itemset x at time $t2$ s.t. Membership of x at time $t2$, $M_x^{t2} = P_j$, where P_i and P_j denote two partitions of the support space, s.t. $t1 < t2$ and $i < j$, then a positive leap, P_{Leap} is defined as:

$$P_{\text{Leap}} = (j - i == 2)?1 : 0$$

and a negative leap, N_{Leap} is defined as:

$$N_{\text{Leap}} = (j - i == -2)?1 : 0$$

Notice that the membership function and the metrics P_{Hop}, N_{Hop}, P_{leap}, N_{leap} represent restructured form of support and therefore represent knowledge at abstraction level zero.

4.2.3 Abstraction

Abstraction is realized at two levels in the proposed change model, using appropriate consolidation functions, defined in Section 4.2.4. Each abstraction level L is associated with suitably defined consolidation functions that yield change metrics of order L+1. Accordingly, Order I and Order II change metrics are defined in the proposed change model.

The Order I change metrics abstract the changes in support of an itemset over smaller periods of time. As shown by FIGURE 5, they are derived by consolidating detected changes over a window of smaller time granularity, called first order window, abbreviated as f-window.

Order II change metrics are statements of support changes in itemsets over longer periods of time. They are derived by consolidating detected changes over a larger window, called second order window, abbreviated as s-window, that encompasses several f-windows, as depicted by FIGURE 5. This facilitates a novel, higher level, restructured view of the changes that aids in their comprehension and deployment.

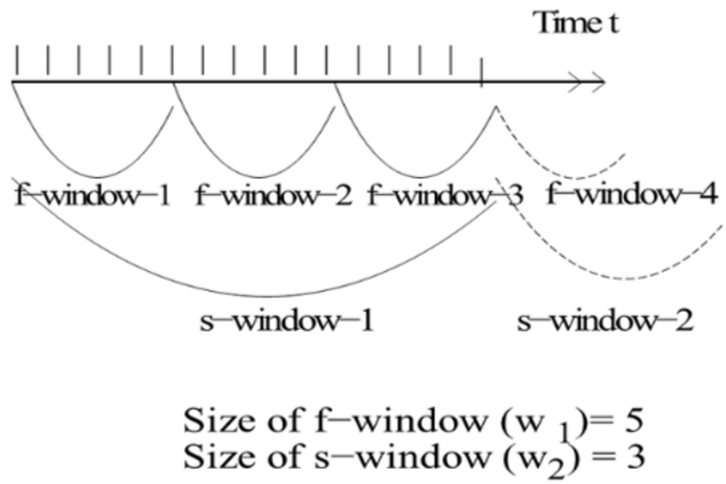

Figure 5: Relation between f-window (for Order I changes) and s-window (for Order II changes).

4.2.4 Consolidation

The design of consolidation functions must be incremental in nature. This follows from the requirement of achieving a higher level of abstraction from a lower level of abstraction, as discussed in Section 4.2.3. Simplicity is another crucial aspect to consider when developing consolidation functions, given that data mining tasks are inherently computationally demanding. In our approach, we rely on additivity as the foundation for consolidation, which offers the advantages of being both incremental and computationally cost-effective.

4.3 Change Metrics of Order I

Two order I metrics, namely, $\text{Jump}_{\text{Count}}$ and Jump_{Mag} are defined to consolidate the significant support transitions of an itemset in terms of the hops and leaps it has made over the smaller time granularity, first order f-window. In other words, these metrics capture the volatility of the frequency of an itemset at abstraction level 1.

$\text{Jump}_{\text{Count}}$ is defined as an order I change metric which is derived from the Hop and Leap metrics, defined at abstraction level 0. It is designed to capture the number of jumps, exhibited by the support behavior of an itemset and is defined as the sum of positive and negative hops and leaps.

$$Jump_{\text{Count}} = (\#P_{\text{Hop}} + \#N_{\text{Hop}} + \#P_{\text{leap}} + \#N_{\text{leap}})$$

Jump_{Mag} is defined as an order I change metric, which is also derived from the Hop and Leap metrics. It is designed to capture the magnitude of the jump in the support behavior of an itemset. It is derived from the weighted aggregate of all positive and negative hops and leaps and is defined as:

$$Jump_{\text{Mag}} = \alpha(\#P_{\text{Hop}} - \#N_{\text{Hop}}) + \beta(\#P_{\text{leap}} - \#N_{\text{leap}})$$

Since, by definition, Hops and Leaps are meant to denote small and big support changes respectively, they are associated with user configurable weights α and β. These weights will be proportionate to the number of partition changes exhibited by the support of an itemset. For instance, in the given framework with three partitions, α associated with Hops is taken to be one, since a hop represents one partition change, while β associated with Leaps is taken to be two, since a leap represents two partition changes in the support of an itemset. Intuitively, these weights are necessary to ensure that the hitherto binary variables Hops and Leaps are quantified proportionate to the underlying support change in itemsets.

4.4 Change Metrics of Order II

Evolution is an expression of the consolidation of changes exhibited by the support of an itemset, at abstraction level II. It is defined as an order II change metric, which is derived from $\mathrm{Jump_{Mag}}$, the change metric of order I. Average is used as the consolidation function and then categorization is done based on the user specified thresholds.

Evolution of an itemset x, $\mathrm{E_x}$ is a mapping from the closed interval [0,1] to the categories Progressive, Stable, Regressive over an s-window, defined as follows:

$$E_{\mathrm{x}} = \begin{Bmatrix} Progressive\ Avg(Jump_{\mathrm{Mag}}) \geq \theta_{\mathrm{progressive}}, \\ Stable\theta_{\mathrm{stable}} \leq Avg(Jump_{\mathrm{Mag}}) < \theta_{\mathrm{progressive}}, \\ Regressive\ Avg(Jump_{\mathrm{Mag}}) < \theta_{\mathrm{stable}} \end{Bmatrix}$$

where $\theta_{\mathrm{progressive}}$ and θ_{stable} are the cut points specified by the user such that $0 \leq \theta_{\mathrm{stable}} \leq \theta_{\mathrm{progressive}} \leq p$, where p is the number of partitions. It's important to highlight that while the computation of Evolution is done in an objective manner, its mapping incorporates user subjectivity.

5. IMPLEMENTATION

The framework presented in this paper was implemented using python 3.6.9., and the Google Colaboratory. Approximately 41.87 GB of disk space and 1.18 GB of RAM were used. Twitter was chosen to source the data for the framework. During the pandemic, when medical facilities in many nations were at a standstill, approaching medical professionals for concerns about women's menstrual health definitely came second to many [9]. In such a scenario, many women found refuge by tweeting their problems, and soliciting the opinions of others on the microblog service Twitter. English tweets regard ing COVID-19 and menses, monthly scraped from the social media site, during the period of 01 January 2021 - 31 December 2021 (FIGURE 6) were used for the study.

5.1 Monitoring Space

Monitoring Space: The support thresholds $S_{\mathrm{h}} = 0.4$, $S_{\mathrm{l}} = 0.2$, were specified to partition the support space into three partitions, corresponding to the infrequent itemsets, sub frequent itemsets, and the frequent itemsets. The aforementioned support thresholds were selected after experimenting with three sets of support thresholds. When high support thresholds were tested, it became evident

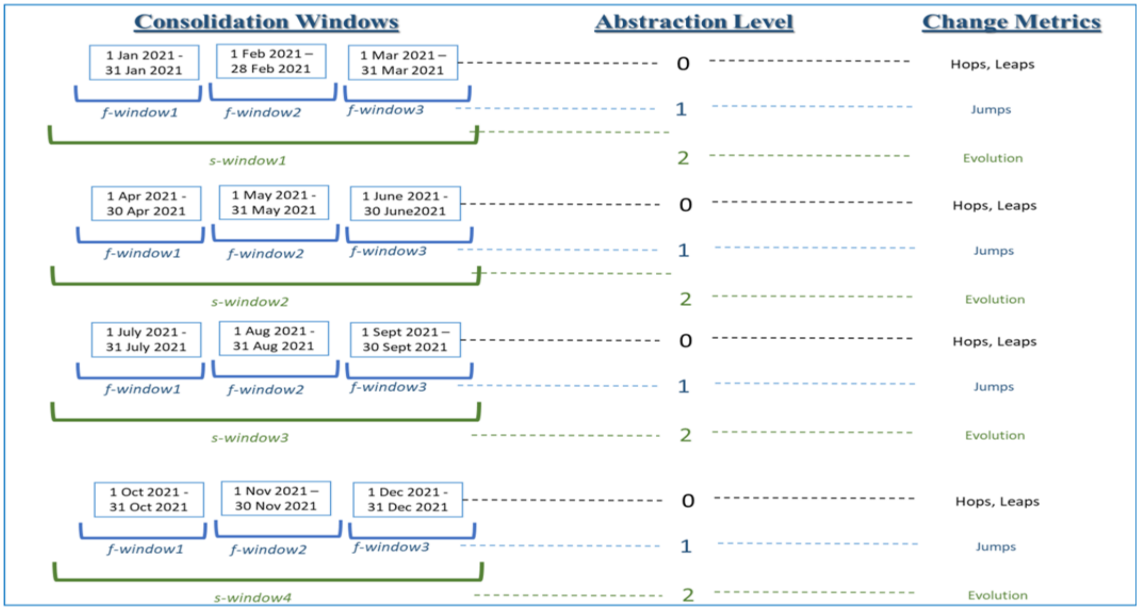

Figure 6: Details of the implementation of the framework.

that the majority of itemsets did not exhibit sufficiently high support levels. Therefore, lower support thresholds were adopted to closely track the support behavior of most itemsets.

Monitoring Itemsets: Itemsets from the MYTH and FACT classes were chosen to be monitored. In all, 262 itemsets were generated from the tweets of MYTH class and 60 itemsets were selected from the FACT class. A total of 15272 rules were discovered from these two classes of itemsets when the itemsets were initially not filtered out for study [10]. As previously mentioned, to comprehend such an extensive output, a significant amount of human time and effort is necessary. It is for this purpose that post-processing filters were proposed in the Knowledge differentiation framework presented by Kochhar et al. [10]. In this framework we take a step ahead and restrict the number of items at the very beginning. Accordingly, the itemsets from only two categories viz. MYTH and FACT that were selected to be monitored for this study are presented in TABLE 1.

It is pertinent to mention here we picked up the largest itemsets for study, though their subsets may have been initially flagged off as 'interesting for study', in order to curtail the redundant association rules and inferences.

5.2 Change Detection, Abstraction, and Consolidation

Changes in support of the itemsets were detected by tracking Hops and Leaps. The consolidation windows were chosen as follows:

1. Size of the first order window i,e, f-window = 1 month. Change metrics of order L, at the abstraction level L = 1, derived from f-window: $Jump_{Count}$ and $Jump_{Mag}$.

Table 1: Monitoring Itemsets

S.No.	Itemset	Itemset Size	Itemset Category
1	{Covid, Stress}	2	FACT
2	{Irregular, COVID}	2	FACT
3	{Pain, Menstrual, COVID}	3	FACT
4	{Heavy, Menstrual, COVID}	3	FACT
5	{Vaccine, Periods, Women}	3	FACT
1	{Vaccinated, Long, Effects}	3	MYTH
2	{Evidence, Vaccine, Fertility}	3	MYTH
3	{May, COVID, Infertility, Cause, Vaccine}	5	MYTH
4	{COVID, Ivermectin, Fertility}	3	MYTH
5	{Women, Menstrual, Medical, Changes, Experts, Dismissing, Step,COVID}	8	MYTH

2. Size of the second order window i,e, s-window = 3 f-windows = 3 months. Change metric of order I, at the abstraction level L = 2, derived from s-window: Evolution.

5.3 Association Rule Generation

Using the above framework, association rules can be generated only from the itemsets whose support behavior and category seems useful to the user. The details association rule generation using the Apriori algorithm [28], are already laid down in the extended framework presented in [10]. It is notable that a few selected, well-categorized rules will be generated from the framework.

6. RESULTS AND INFERENCES

This section presents the results of the implemented framework and the inferences drawn.

6.1 Inferences for FACT Itemsets

A comparison of the positive and negative Hops, positive and negative Leaps, for the set of monitored itemsets M^1 belonging to the category FACT, for a period of three f-windows i.e., 3 months, viz. January 2021 – March 2021, is presented in TABLE 2. The table also shows the computation of the order I metrics $\text{Jump}_{\text{Count}}$ and Jump_{Mag} over the three f-windows. Computation of the order II metric Evolution, with $\theta_{\text{progressive}} = 1$ and $\theta_{\text{stable}} = 0$, over the first s-window, encompassing the first three f-windows is also depicted in TABLE 2.

The following inferences can be drawn from the statistics presented in TABLE 2 for FACT itemsets:

1. The 2-itemset Irregular, COVID indicates that people took to Twitter to complain about irregular periods in relation to the COVID-19 pandemic. The support behaviour of the itemset

Table 2: Change Metrics of Order I and II for FACT itemsets

S.No.	Itemset	Change Metrics						s-window1
		f-window1		**f-window2**		**f-window3**		
1	**{Irregular, COVID}**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	Evolution {Irregular, COVID}
		2	2	0	2	0	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		1	0	0	0	1	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		5	**2**	2	**-2**	1	**2**	Progressive
2		**f-window1**		**f-window2**		**f-window3**		Evolution {Covid, Stress}
	{Covid, Stress}	**#PHop**	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		1	0	1	1	1	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		1	**1**	2	**0**	1	**1**	Progressive
3		**f-window1**		**f-window2**		**f-window3**		Evolution {Pain, Menstrual, COVID}
	{Pain, Menstrual, COVID}	**#PHop**	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	0	1	1	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		1	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		1	**2**	1	**-1**	1	**1**	Progressive
4		**f-window1**		**f-window2**		**f-window3**		Evolution {Heavy, Menstrual, COVID}
	{Heavy, Menstrual, COVID}	**#PHop**	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		1	0	0	0	1	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		1	**1**	0	**0**	1	**1**	Progressive
5		**f-window1**		**f-window2**		**f-window3**		Evolution {Heavy, Menstrual, COVID}
	{Vaccine, Periods, Women}	**#PHop**	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	0	1	0	1	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		1	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		1	**2**	1	**-1**	1	**-1**	Stable

is quite volatile as evident from high values of $\text{Jump}_{\text{Count}}$ in all three f-windows. In fact the f-window1 shows the highest $\text{Jump}_{\text{Count}}$ of all the FACT itemsets monitored. The variation in the support count, depicted by Jump_{Mag} from +2 to -2 and back to +2 over the three f-windows indicates that the support of the itemset increased so much in the first f-window that it jumped two partitions, dropped again by two partitions in the next f-window before coming back to the high support partition with a jump of +2 in Jump_{Mag}. This further strengthens the proof of volatility of the support of itemset.

The overall average Jump_{Mag} (=2) is greater than $\theta_{\text{progressive}} = 1$, and indicates that on an average the support of itemset has exhibited a positive transition that enabled its jump of greater than or equal to one partition. Thus the itemset shows an overall progressive support behavior in the s-window i.e. first quarter of 2021, which is faithfully captured by the change metric of order II, Evolution.

2. Though the 2-itemset Covid, Stress also exhibits average Jump_{Mag} (=2), greater than $\theta_{\text{progressive}} = 1$, and its evolution status is labeled Progressive, its $\text{Jump}_{\text{Count}}$ in the three f-windows are lower. This indicates lesser number of times the variation in its support is noticed. Lower number of hops and leaps than the itemset Irregular, COVID provide proof of this inference.

3. The support behaviour of itemset Pain, Menstrual, COVID is identical to the itemset Covid, Stress, with average magnitude of support i.e. Jump_{Mag} =2, and evolution status Progressive. The count of variation in its support is also low as indicated by $\text{Jump}_{\text{Count}}$.

4. Significant positive transitions in support behaviour of itemset Heavy, Menstrual, COVID can be observed in the first and third f-windows, leading to a $\text{Jump}_{\text{Count}} = 1$ and Jump_{Mag} =1 in both these f-windows. It is noteworthy that though even for this itemset the average Jump_{Mag} =2, and evolution status is Progressive, a look at the change metrics such as Hop and Leap provide insights of exactly how and when the support behaviour of the itemset changed over the quarter of 2021.

5. A look at the Hop and Leap metrics for the Vaccine, Periods, Women indicates that its support took a leap in the first month i.e. Jan 2021, possibly the time around which the COVID vaccines were being introduced in most countries around the world. The anxiety about the possible adverse effects of a new vaccination and therefore the talk on Twitter at this time was high, leading to the positive Leap in support of the itemset. The next two months see a negative Hop, and a positive Hop, indicating a decline in the talk in February 2021 and then again some buzz on the topic in March 2021. The overall $\text{Jump}_{\text{Mag}} = 1$, indicating an overall jump of one support partition, bringing the Evolution status to Stable.

6.2 Inferences for MYTH Itemsets

TABLE 3 shows the order I and order II change metrics computed for the same quarter as the FACT itemsets i.e., Jan-March 2021. As earlier, the order I metrics were computed over one s-window, encompassing three f-windows of one month each.

The following inferences can be drawn from the change metrics computed on MYTH itemsets (TABLE 3):

Table 3: Change Metrics of Order I and II for MYTH itemsets

1	**{Vaccinated, Long, Effects}**	**f-window1**		**f-window2**		**f-window3**		Evolution {Vaccinated, Long, Effects}
		#PHop	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	1	0	0	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		0	**0**	1	**1**	0	**0**	Progressive
2	**{Evidence, Vaccine, menses}**	**f-window1**		**f-window2**		**f-window3**		Evolution {Evidence, Vaccine, menses}
		#PHop	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		1	0	0	0	0	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	1	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		1	**1**	0	**0**	1	**-2**	Regressive
3	**{May, COVID, Infertility, Cause, Vaccine}**	**f-window1**		**f-window2**		**f-window3**		Evolution {May, COVID, Infertility, Cause, Vaccine}
		#PHop	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	1	0	0	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		0	**0**	1	**1**	0	**0**	Progressive
4	**{COVID, Ivermectin, Fertility}**	**f-window1**		**f-window2**		**f-window3**		Evolution {COVID, Ivermectin, Fertility}
		#PHop	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	0	0	1	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	0	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		0	**0**	0	**0**	1	**1**	Progressive
5	**{Women, Menstrual, Medical, Changes, Experts, Dismissing, Step, COVID}**	**f-window1**		**f-window2**		**f-window3**		Evolution {Women, Menstrual, Medical, Changes, Experts, Dismissing, Step, COVID}
		#PHop	**#NHop**	**#PHop**	**#NHop**	**#PHop**	**#NHop**	
		0	0	0	0	0	0	
		#Pleap	**#Nleap**	**#Pleap**	**#Nleap**	**#Pleap**	**#Nleap**	
		0	0	0	0	1	0	
		Jump Count	**Jump Mag**	**Jump Count**	**Jump Mag**	**Jump Count**	**Jump Mag**	
		0	**0**	0	**0**	1	**2**	Progressive

1. The itemset Vaccinated, Long, Effects suggests that there was a belief among people that a vaccinated population would experience some long-term effects. Given that the context of the discussions was focused on menstrual disorders, we can further refine the inference to conclude that vaccines might have enduring impacts on the menstrual cycle. However, a thorough examination of medical papers and news articles has subsequently debunked this rule as a MYTH, as it was found that, in most cases where the vaccine did affect the menstrual cycle, these effects were only temporary [27].

   The itemset, which was initially placed in the low support partition doesn't show any significant transition, which is captured as zero Hops and Leaps in all three f-windows. The overall Jump_{Mag} was therefore zero, and the Evolution status was Stable (in its original support partition).

2. The itemset Evidence, Vaccine, menses, flagged off as MYTH, suggests a proof of the effect of COVID-19 vaccination on menses. However, clearly being a rumor, the itemset initially saw an increase in talk in the first f-window, around the time of introduction of vaccines, and then saw a stark decrease in support in the last f-window, bringing the overall $\text{Jump}_{\text{Mag}} = -1$, and the Evolution status to Regressive.

3. Another interesting discovered against the backdrop of menstrual conversations, May, COVID, Infertility, Cause, Vaccine, suggested that there was supporting evidence indicating that COVID-19 vaccines could hinder the likelihood of pregnancy. However, our subsequent investigation revealed that individuals had misconstrued the spike protein responsible for placental growth and formation with the spike protein found on the surface of the COVID-19 virus. In reality, the COVID-19 vaccine targets the latter spike protein, and confusion arose because people mistakenly believed that since the vaccine affects this protein, it might also interfere with placental formation, potentially causing infertility [29].

   The itemset was detected in partition 1 i.e. the low support group. A positive Hop in the second f-window led to the overall magnitude of Jump for the itemset as one and the Evolution status as Progressive.

4. The itemset COVID, Ivermectin, Fertility largely pointed to the similar conversation as the itemset May, COVID, Infertility, Cause, Vaccine, with the exception that it referred to a particular medication, Ivermectin, which was not originally designed as a COVID-19 vaccine but rather utilized in its therapy. Through further inquiry, it was ascertained that certain individuals disseminated false information regarding the potential infertility caused by Ivermectin, relying on a study conducted in 2019 [30]. This itemset was also detected in partition 1, i.e., the low support group. A positive Hop in the third f-window led to the overall magnitude of Jump for the itemset as one and the Evolution status as Progressive.

5. The itemset Women, Menstrual, Medical, Changes, Experts, Dismissing, Step,COVID was a result of the messages shared by individuals on social media platforms, expressing their frustration that COVID experts had disregarded reports of menstrual changes in women. Despite the fact that the menstrual changes observed in women due to COVID infection or vaccination have generally been transient, people were clamoring for a comprehensive examination of any potential long-term effects.

A look at the change metrics for the itemset shows a positive leap in its support in t the third f-window, leading to an overall $\text{Jump}_{\text{Mag}} = -2$, and the Evolution status to Progressive.

6.3 Association Rule Generation

The 'Regressive' category of itemsets was chosen for the generation of Association rules, which comprised of the itemset Evidence, Vaccine, menses. The following strong rule (high confidence of 90%), with positively correlated antecedent and consequent (lift>1), was generated:

It can be observed from TABLE 4 that only a single rule, well labeled with MYTH and regressive categories is generated. Similarly, a limited number of well-labeled rules from the other categories also can be generated, if desired.

Table 4: The Generated Association Rules

Rule LABEL	Antecedent	Consequent	Confidence	lift
MYTH, REGRESSIVE	({'evidence', 'vaccine'})	({'menses'})	0.9	1.45

7. CONCLUSIONS AND FUTURE DIRECTIONS

The work presented in this paper highlights how the pandemic has caused stress that can affect menstrual cycles and emphasizes the need for comprehensive research in this area, both from the medical community and the machine learning community.

Accordingly, a framework that employs change modeling is implemented to track COVID-19-related menstrual health issues from Twitter data over time. This approach aims to provide a consolidated and abstract view of the changes in the support of words that people use on Twitter. Two first-order metrics, $\text{Jump}_{\text{Count}}$ and Jump_{Mag}, have been designed to consolidate the count and amount of significant support transitions of an itemset by observing the number of hops and leaps it has made during a smaller time granularity, first-order window called a f-window. These metrics gauge the volatility in the frequency of an itemset. On the other hand, Evolution represents the aggregation of changes observed in the support of an itemset at the second abstraction level (abstraction level II). It is characterized as a second-order change metric and is derived from Jump_{Mag}, which is a first-order change metric. Additionally, the framework generates a small number of well-labeled association linkages from Twitter discussions on menstruation, focusing on creating understandable and context-aware knowledge insights.

To conclude, the framework produces simple, intuitive, one-look metrics for tracking changes from multiple perspectives, aiding in decision-making, and offers computational advantages through the generation of a reduced number of relevant association rules.

As a future work, an extension of the study to pre-COVID and post-COVID time periods can be explored. Such a study will provide avenues for comparing and understanding the evolution of menstrual health problem patterns over a relatively larger time span. Similarly, some physical factors can introduce confounding variables that can affect the study, for instance, the geographical location and region from which the tweets originated, as well as the severity and frequency of health issues reported in each specific location. The interplay and effect of such factors can be explored in the next work.

8. CONFLICT OF INTEREST

The authors declare that they have no conflicting interests.

References

[1] Ozimek N, Velez K, Anvari H, Butler L, Goldman KN, et al. Impact of Stress on Menstrual Cyclicity During the Coronavirus Disease 2019 Pandemic: A Survey Study. J Womens Health (Larchmt). 2022;31:84-90.

[2] Plotsky PM, Owens MJ, Nemeroff CB. Psychoneuroendocrinology of Depression. Hypothalamic-Pituitary-Adrenal Axis. Psychiatr Clin North Am. 1998;21:293-307.

[3] Yunitri N, Chu H, Kang XL, Jen HJ, Pien LC, et al. Global Prevalence and Associated Risk Factors of Posttraumatic Stress Disorder During COVID-19 Pandemic: A Meta-Analysis. Int J Nurs Stud. 2022;126:104136.

[4] Demir O, Sal H, Comba C. Triangle of COVID, Anxiety, and Menstrual Cycle. J Obstet Gynaecol. 2021;41:1257-1261.

[5] Wang YX, Arvizu M, Rich-Edwards JW, Stuart JJ, Manson JE, et al. Menstrual Cycle Regularity and Length Across the Reproductive Lifespan and Risk of Premature Mortality: Prospective Cohort Study. BMJ. 2020;371:m3464.

[6] Wang YX, Stuart JJ, Rich-Edwards JW, Missmer SA, Rexrode KM, et al. Menstrual Cycle Regularity and Length Across the Reproductive Lifespan and Risk of Cardiovascular Disease. JAMA Netw Open. 2022;5:e2238513.

[7] Alvergne A. Why We Must Fight Ignorance About COVID-19 Vaccines and Menstrual Cycles. Trends Mol Med. 2023;29:678-680.

[8] Lebar V, Laganà AS, Chiantera V, Kunič T, Lukanović D, et al. The Effect of COVID-19 on the Menstrual Cycle: A Systematic Review. J Clin Med. 2022;11:3800.

[9] Muhaidat N, Alshrouf MA, Azzam MI, Karam AM, Al-Nazer MW, et al. Menstrual Symptoms After COVID-19 Vaccine: A Cross-Sectional Investigation in the Mena Region. Int J Womens Health. 2022;14:395-404.

[10] Kochhar SK, Kaur R. Breaking the Taboos: Deploying Knowledge Differentiation to Study COVID-19 Ramifications on Women's Menstrual Health. Int J Comput Digit Syst. 2023;13:1.

[11] Cao L. Actionable Knowledge Discovery and Delivery. WIREs Data Min Knowl Discov. 2012;2:149-163.

[12] Cao L, Luo D, Zhang C. Knowledge Actionability: Satisfying Technical and Business Interestingness. Int J Bus Intell Data Min. 2007;2:496-514.

[13] Kalanat N. An Overview of Actionable Knowledge Discovery Techniques. J Intell Inf Syst. 2022;58:591-611.

[14] Nazir M, Asghar S, Rathore MA, Shahzad A, Shahid A, et al. Menstrual Abnormalities After COVID-19 Vaccines: A Systematic Review. Vacunas. 2022;23:S77-S87.

[15] Chao MJ, Menon C, Elgendi M. Effect of COVID-19 Vaccination on the Menstrual Cycle. Front Med. 2022;9.

[16] Baena-García L, Aparicio VA, Molina-López A, Aranda P, Cámara-Roca L, et al. Premenstrual and Menstrual Changes Reported After COVID-19 Vaccination: The Eva Project. Womens Health (Lond). 2022;18.

[17] Bhatnagar V, Kochhar SK. Modeling Support Changes in Streaming Item Sets. Int J Syst Sci. 2006;37:879-891.

[18] Bhatnagar V, Kochhar SK. User Subjectivity in Change Modeling of Streaming Itemsets. In: Li X, Wang S, Dong ZY, editors. Advanced data mining and applications. Heidelberg: Berlin. Berlin, Heidelberg: Springer. 2005:812-823.

[19] Haque A, Khan L, Baron M, Thuraisingham B, Aggarwal C. Efficient Handling of Concept Drift and Concept Evolution Over Stream Data. In: 32nd International Conference on Data Engineering (ICDE). IEEE Publications. 2016;481-492.

[20] Gupta A, Bhatnagar V, Kumar N. Mining Closed Itemsets in Data Stream Using Formal Concept Analysis. In: Pedersen TB, Mohania MK, Tjoa AM, editors. Data warehousing and knowledge discovery. Berlin, Heidelberg: Springer. 2010:285-296.

[21] Gupta A, Kumar N, Bhatnagar V. Incremental Classification Rules Based on Association Rules Using Formal Concept Analysis. In: Perner P, Imiya A. (eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2005. Lecture Notes in Computer Science,Springer, Berlin, Heidelberg. 2005;3587.

[22] Gupta A, Kumar N, Bhatnagar V. Mining of Multiobjective Non-redundant Association Rules in Data Streams. In: International Conference on Artificial Intelligence and Soft Computing. 2012;73-81.

[23] Kumar N, Gupta A, Bhatnagar V. Lattice Based Associative Classifier. In Pan JS, Chen SM, Nguyen NT, editors. Intelligent information and database systems. Heidelberg: Berlin. Berlin, Heidelberg: Springer. 2012:22-31.

[24] Cleland ZW, Dao KA, Dao THD. Detecting Changes in Spatial Characteristics of Colorado Human-Caused Wildfires Using Apriori-Based Frequent Itemset Mining. Comput Environ Urban Syst. 2023;101:101941.

[25] Yin Y, Li P, Chen J. A Variable Sliding Window Algorithm Based on Concept Drift for Frequent Pattern Mining Over Data Streams. In IEEE 28th International Conference on Parallel and Distributed Systems (ICPADS). 2023;818-825.

[26] Swetanisha S, Panda AR, Behera DK. Change Detection Using Machine Learning Models: A Case Study on the Puri District of Odisha, India. In: 19th OITS International Conference on Information Technology (OCIT). 2021;100-104.

[27] Liu B, Hsu W, Ma Y. Mining Association Rules With Multiple Minimum Supports. In: Proceedings of the fifth ACM SIGKDD international conference on knowledge discovery and data mining, KDD'99. NY, USA. Association for Computing Machinery. 1999;337-441.

[28] Agrawal R, Srikant R. Fast Algorithms for Mining Association Rules. In Proc 20th Int Conf. Very Large Data Bases VLDB. 1994;12:487-499.

[29] Taylor S. COVID stress syndrome: Clinical and Nosological Considerations. Curr Psychiatry Rep. 2021;23:1-7.

[30] Mekuriaw W, Balkew M, Messenger LA, Yewhalaw D, Woyessa A, et al. The Effect of Ivermectin® on Fertility, Fecundity and Mortality of Anopheles Arabiensis Fed on Treated Men in Ethiopia. Malar J 2019;18:357.