Developing Human-centric Machine Learning Models for Temporal Data

Author information scrubbed for double-blind reviewing

No Institute Given

Abstract. Involving humans at every stage of developing a machine learning model is crucial for making AI systems more human-centric, both in model development and generating explanations. In this work, we developed an approach to building and iteratively improving a machine learning model with involvement of human-gained knowledge using a Spanish COVID-19 dataset as a test bed. This approach was then generalized for application to other data describing temporal phenomena, processes, or events. The proposed method utilized human insights obtained through visual analytics techniques applied to the data and the model output. By incorporating these human-gained insights into the model, performance improved and a greater understanding of the relationships between the data attributes was achieved. The insights from the COVID-19 case study were used to propose a generic workflow for developing human-centric models for temporal data. Additionally, the knowledge gained from the modeling process can potentially be used for the generation of human-centric explanations.

Keywords: Human-centric modelling · Temporal Modelling · Visual Analytics

1 Introduction

Machine learning (ML) models are often built in a purely data-driven manner, selecting features that optimize performance metrics such as mean squared error or area under the ROC curve. However, these approaches have limitations that can be mitigated by incorporating domain expert knowledge. Expert knowledge can help derive more expressive features than the original data variables and create meaningful feature combinations aligned with human mental models, enhancing the interpretability and trustworthiness of the resulting ML models. While numeric performance indicators do not provide insights into model errors or guide improvement efforts, human experts can analyze errors in detail and find ways to improve based on their domain understanding. Additionally, purely data-driven models can struggle with overfitting or underfitting, particularly with noisy or insufficient data. Human guidance can address data quality issues, ensuring better generalization.

Human-driven model building fosters a collaborative approach where domain knowledge and machine learning techniques complement each other. This collaboration can lead to solutions that neither approach could achieve independently.

Experts can propose hypotheses and test them using ML models, iteratively refining both the model and their understanding of the domain. In dynamic environments, such as the COVID-19 pandemic, where underlying patterns and relationships can evolve rapidly, human experts can continuously update the model with new insights and data, ensuring it remains relevant and accurate.

Our research focuses on models dealing with sequences of events that reflect the evolution or progression of temporal phenomena through identifiable stages or manifestations. The modeling task involves representing the interplay of two or more phenomena that may influence each other. We develop our approach within a case study on modeling the interrelationships between COVID-19 dynamics and changes in population mobility behavior. The goal of the model is to predict the next pandemic event based on preceding events and mobility behaviors.

We build and iteratively improve an ML model using visualizations to understand the phenomena and model properties. We incorporate human insights into the model through data operations rather than specialized algorithms, aiming for compatibility with common ML techniques. This empirical study is synthesized into a general framework specifying visualizations, data operations, and interaction techniques to support human-driven model development for temporal data. This paper presents our case study, followed by the general framework derived from our experiences, focusing on human-driven model development for temporal phenomena.

2 Related Work

In this work, we propose a methodology for model building that leverages human knowledge obtained through visual analytics (VA) for temporal data. We first introduce the current research landscape in human-centered modeling and some aspects of temporal modeling.

The incorporation of prior knowledge into the ML process is known as "informed ML". Unlike traditional ML methods using prior knowledge indirectly through feature engineering or hyperparameter tuning, informed ML utilizes knowledge represented explicitly in machine-processable forms, such as algebraic equations, logic rules, or knowledge graphs [19, 5]. While VA can potentially support the acquisition of knowledge from human experts, including both prior knowledge and newly obtained insights from visual analysis [2], we are not aware of existing VA systems that facilitate externalization of expert knowledge for ML. However, there are instances of using VA for developing conceptual models, even if not directly intended for ML integration [7, 23].

More traditionally, VA can facilitate involvement of human knowledge through preparation of input data for ML (selection, cleaning, feature engineering, etc.) and iterative model refinement. Liu *et al.* [14] proposed a VA framework for steering data quality. Hong *et al.* developed a VA system that supports imputation and transformation techniques to improve data quality, enhancing ML performance [9]. The system also aids in visual interactive labelling and other steps of data preparation. Andrienko *et al.* propose a VA approach for generating descriptive features of movement patterns and interactive labeling for classifier training [3]. Muhlbacher and Piringer describe using VA to iteratively refine regression models [18]. Comprehensive surveys highlight VA's role in data science workflows and ML model building [16, 22, 21].

Human-AI interaction techniques support attaining the goals of explainable and interactive ML, which include understanding and justifying model decisions, diagnosing model failures, and refining models to fit application domains [15, 20]. VA frameworks like "explAIner" [20] combine interactive ML and VA for conducting a dialog with a user in which explanations are iteratively refined.

For temporal data, human knowledge is integrated in ML models through converting raw data into interpretable patterns represented by combinations of descriptive features. This operation is called "temporal abstraction" [4]. From a large variety of summary statistics that can be derived from time series data [17], selecting relevant and understandable features remains challenging. Andrienko *et al.* [3] propose a VA approach transforming time series to episodes which capture the patterns of interest and computing key attributes to form training data for pattern classification. However, this focuses on data preparation without addressing model development. Our work extends this by using temporal abstraction and pattern classification for predictive modeling and applying VA techniques throughout the model development process, including evaluation, comparison, and iterative refinement.

3 Case Study

As stated in the introduction, our task is to predict the class of the next COVID-19 event (i.e., the level of morbidity) based on preceding disease events and levels of population mobility. The data consist of two parallel event sequences: one for disease events and another for mobility events. These parallel sequences are available for different geographic areas; in our case study, these are 52 provinces of Spain. The original data were provided in the form of continuous numeric time series. We transformed them to event sequences, where each event spans over a time interval in which the class (level) of disease morbidity or mobility behavior is distinct from those in the preceding and following intervals. While the two event sequences unfold in parallel over time, there is no synchronization between the events, and they have variable duration.

3.1 Original data

We used two openly available datasets prepared by Ponce-de-Leon et al. [10]: daily counts of new COVID-19 cases for the 52 provinces of Spain covering the period from 01/01/2020 to 12/08/2021 [11] and daily counts of trips within and between the provinces covering the period from 14/02/2020 to 09/05/2021 [12]. For our analysis and modeling, we used data from 17/02/2020 (Monday) to 09/05/2021 (Sunday), encompassing 64 full weeks. We transformed the disease case counts into counts per 100,000 province inhabitants using openly available population data [13]. The counts of trips within the provinces were normalized to the pre-pandemic levels still attained in the period from 14/02/2020 to 10/03/2020, as can be seen in Fig.1, bottom.

Based on general knowledge about the impacts of people's mobility on the spread of contagious diseases and specific historical knowledge about the pandemic and mobility reduction measures taken in different countries, we hypothesized that the morbidity and mobility time series are linked by two-way relationships. Changes in disease levels may lead to adaptations in population mobility behaviors, and changes in mobility behaviors affect the further evolution of the pandemic. However, the effects of each phenomenon on the other become noticeable after certain time lags. To determine the temporal horizons of these effects, we conducted cross-correlation and Granger causality [8] analyses. Visual displays, as shown in Fig.2, were used to explore the correlation and causality coefficients across different time lags. The calculations were done separately for three time periods: 17/02/2020 to 30/04/2020; 01/08/2020 to 15/11/2020; and 01/01/2021 to 09/05/2021. These periods roughly correspond to three major waves in the pandemic evolution visible in the upper graph in Fig.1. We considered these periods separately because population mobility behaviors did not change uniformly across them, as seen in the lower graph in Fig.1.

The graphs in Fig.2 show that the three periods indeed exhibited different relationships between the pandemic and mobility levels. Unlike the first period, the second and third periods are characterized by significant variation in the cross-impacts between the two phenomena across the provinces. Nevertheless, the first period, marked by large and consistent changes in mobility across provinces, allows us to determine the time ranges in which changes in one phenomenon may affect the other. Specifically, we found that mobility affects morbidity levels within 14 to 35 days, and morbidity levels affect mobility within 7 to 14 days. These findings guided the further analysis.

3.2 Data processing

To minimize the influence of the irrelevant weekly cycle in the data variation, we smoothed both sets of time series using a forward moving average with a sliding window of 7 days. We segmented the series into weekly episodes and categorized them into four levels of pandemic severity and four levels of mobility, using an interactive approach with feature-based spatial embedding, clustering, and kNN classification [3] For this purpose, the value variation in the episodes was represented by summary features including the median, maximum, and trend line angle. The COVID-19 levels were denoted as 'c1', 'c2', 'c3', and 'c4' and the mobility levels were labeled as 'm1', 'm2', 'm3', and 'm4', where 'm4' represents normal or near-normal mobility, and 'm1' indicates predominant staying at home. Consecutive episodes of the same category were combined into continuous events of varying duration, creating non-synchronized event sequences.

To visualize the event sequences, we created timeline displays (Fig.3) showing the distributions of event categories over time and across provinces. However, this data format is not directly suitable for ML algorithms. For a model predicting

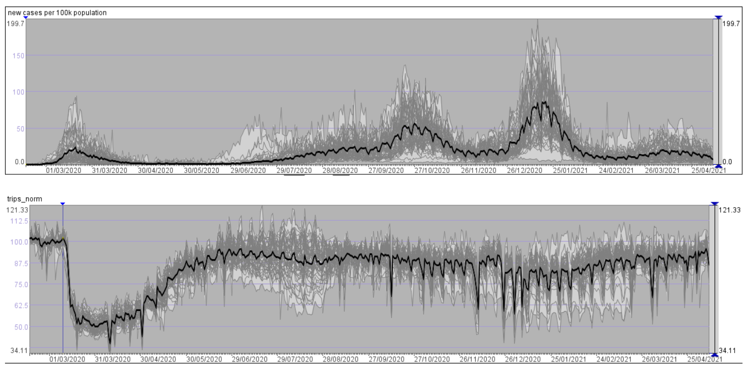


Fig. 1: Upper image: Time series of new COVID-19 cases per 100,000 inhabitants in the Spanish provinces. Lower image: Time series of daily trip counts within provinces, normalized as percentages of the respective values during the normal mobility period from 17/02/2020 to 10/03/2020. The vertical line marks the date 10/03/2020. In both graphs, each grey line represents one province, and the thick black line shows the variation of the country's mean.

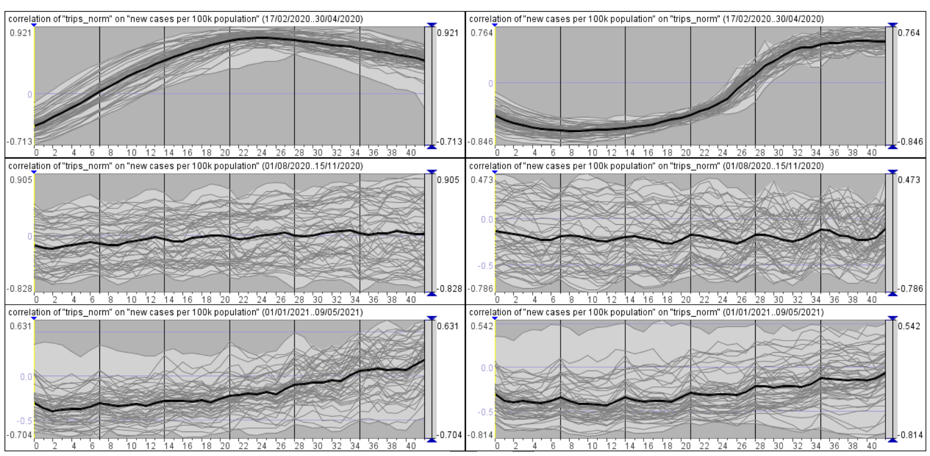


Fig. 2: Series of correlation coefficients between the COVID-19 and mobility time series, calculated over time lags from 0 to 42 days separately for the time periods 17/02/2020-30/04/2020, 01/08/2020-15/11/2020, and 01/01/2021-09/05/2021. The graphs on the left display the lagged correlations of new disease cases to mobility levels, and the graphs on the right show the inverse correlations. The thick black lines connect the mean values over the sequences of time lags.

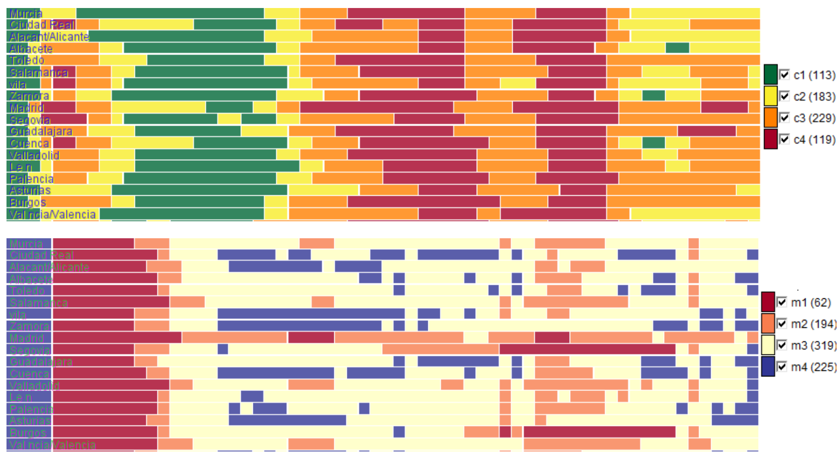


Fig. 3: Fragments of the timeline displays of the pandemic events (top) and mobility events (bottom). The horizontal dimension represents time, while each row corresponds to one province. Events are depicted as horizontal bars positioned according to their occurrence times, with bar lengths proportional to event durations. The colors indicate the event categories; see the legends on the right.

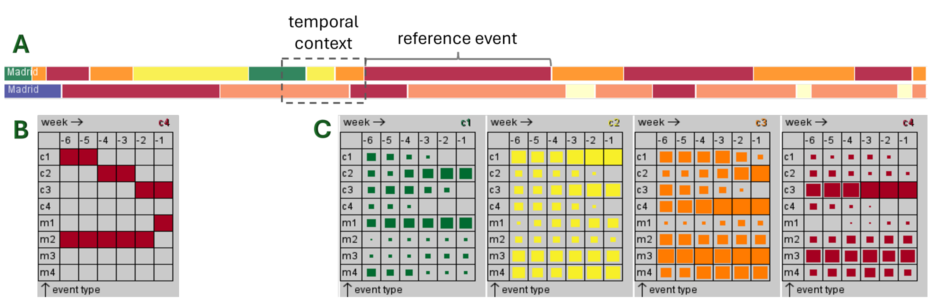


Fig. 4: Extracting and representing temporal contexts of events. A: The context of one reference event. B: The context represented as a binary matrix. C: Aggregated visual representation of temporal contexts of multiple events.

the class of a target pandemic event based on preceding disease and mobility events, we need to represent the preceding events of each pandemic event in the data as features of that event.

We achieve this by extracting the temporal contexts of the pandemic events. The temporal context includes the disease and mobility events within a specified time window before the target event. This is illustrated in Fig.4A using an example reference event. The temporal context is represented by a binary matrix, with columns corresponding to equal-length intervals within the time window and rows representing event types. Values of 1 and 0 indicate the presence or absence of each event type in each interval. In Fig.4B, the values 1 and 0 are visually represented by filled or empty cells. The time window length is 6 weeks, chosen based on the time horizons of possible effects between events (see Section 3.1). It is divided into 6 weekly intervals.

The contexts of individual events can be aggregated by event groups and visualized as shown in Fig.4C. Each matrix corresponds to one class of pandemic events, indicated by the color of the squares in the cells. The square sizes are proportional to the counts of the value 1 in the corresponding positions of individual event matrices.

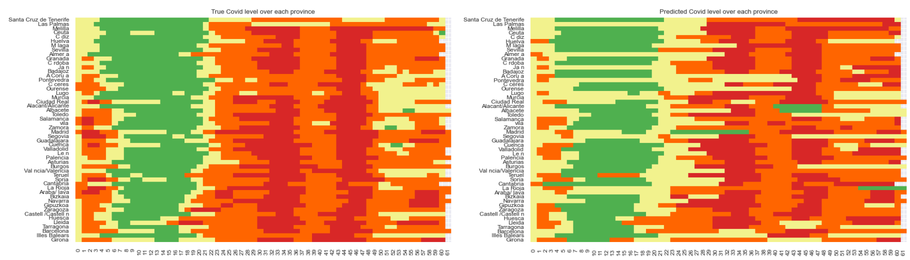
As the categories of both COVID-19 and mobility events are ordered, we can represent the event context alternatively by specifying the ordinal numbers of event categories in each interval, i.e., using numbers from 1 (lowest level) to 4 (highest level). This representation can be more suitable for ML algorithms requiring numeric features. For training and testing our model, we use the transformed data, which include the class (severity level) of each pandemic event and the temporal context from the preceding 6 weeks.

3.3 Model development

Initially, a base model was developed using six-week event contexts to predict the current level of COVID-19 based on historical behavior. To make the model more realistic, we excluded the immediately prior two weeks of pandemic data and one week of mobility data, as such up-to-date data may not be available in practice. We also removed the sixth week prior to the event due to its minimal contribution to predictions.

Initial data exploration suggested that the distinct characteristics of the first pandemic wave might require excluding this period from the model. We also noted geographic differences, specifically disparities between islands and mainland provinces. Nonetheless, we created the first model variant using all data to understand the effects of these temporal and spatial disparities on predictions. For modeling, we employed a random forest algorithm with a set random seed for reproducibility and used a 5-fold cross-validation strategy.

Predictions from the initial base model were visualized in a timeline display similar to Fig.3, top, to observe spatial and temporal patterns. Juxtaposed timeline displays of the true and predicted event classes are shown in Fig.5. The prominent visual differences indicate numerous classification errors. To understand the character and spatio-temporal distribution of these errors, we com-



(a) The true class labels (b) The predicted class labels

Fig. 5: The (a) true versus (b) predicted class labels using the base model for every province plotted along the time period of interest. The class labels are: Green: C1, Yellow: C2, Orange: C3, Red: C4

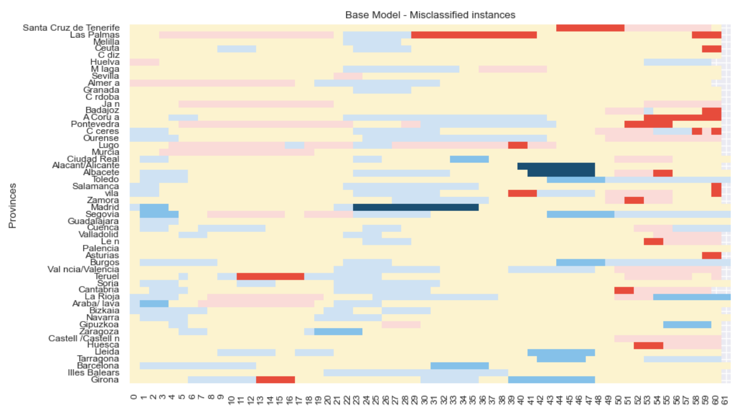


Fig. 6: The prediction errors of the base model. The shades of blue and red represent, respectively, under- and over-estimations, with shade darkness representing the amount of the difference.

puted the differences between the predicted and actual classes, treating categories c1 to c4 as numbers from 1 to 4, and explored the distribution using a timeline display as in Fig.6. It reveals that the model underestimates COVID-19 levels during the first wave. Dependencies learned from this period, apparently, lead to significant underestimations during subsequent waves. This aligns with our anticipation that the distinctiveness of the initial period may have a negative impact on model learning. Therefore, excluding this period from model training is a reasonable strategy.

Additionally, the distribution of model errors confirms that geographic differences, specifically between islands and the mainland, hinder proper generalization and impair model performance. Hence, it makes sense to develop separate

Model	Operations undertaken	5-Fold CV mean accuracy
M1	Basic Model - No extra operations	0.71244
M2	First 30 days removed	0.71827
M3	Number of weeks since the start added as a feature	0.74798
M4	Duration of the event added as a feature	0.78847
M5	All island provinces are removed	0.71636
M6	All of the above operations	0.86334

Table 1: A summary of the operations applied during the iterative model development and the performance of the resulting model versions.

models for the mainland and islands. In our case study, we chose to exclude the islands from consideration.

Based on these observations and examination of subsequent model versions, we iteratively developed six model versions, aiming to improve prediction performance using insights gained from data and model behavior exploration. As mentioned in the introduction (Section 1), we refrained from using or developing special algorithms capable of directly incorporating expert knowledge into the model training process. For common ML algorithms, the only way to utilize expert knowledge is through careful preparation and selection of training data. Hence, we applied a series of modifications to the training dataset, summarized in Table 1, which also displays the accuracy scores of the model versions.

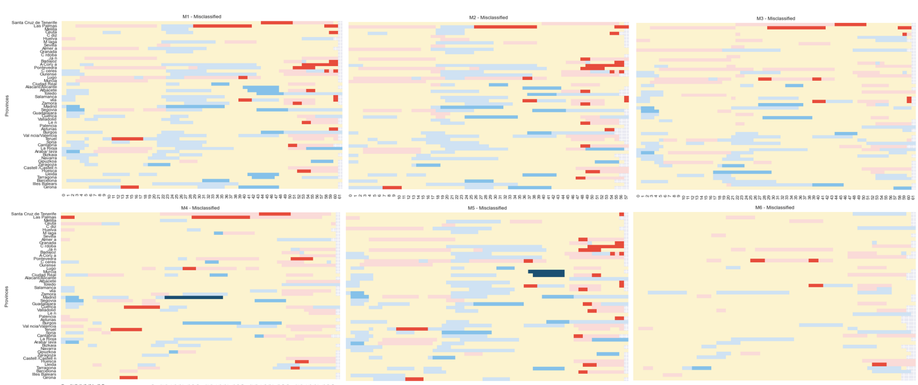


Fig. 7: Error distributions for six model versions. Shades of blue and red signify under- and overestimated events, respectively, with darker shades indicating greater degrees of under- or overestimation.

Figure 7 shows the misclassified events across the six models tested. With all modifications included, the error rate is substantially reduced, and the temporal

Type of knowledge	**Origin**	**Utilization**
Initial period of yet normal mobility	Visualization of the time series of trips counts	Baseline values for transforming the absolute counts to relative
Weekly patterns of mobility and COVID-19 (weekend drops)	Time series plots	Smoothing over 7-days period, use of week-based summary features for defining event categories
Distribution of ranges and trends of weekly variations	2D embedding and clustering of weekly episodes	Categories of COVID-19 and mobility events
Time horizons of possible effects between COVID-19 and mobility	Cross-correlation and Granger causality analyses	Lengths of the temporal contexts of the events to be used for prediction
Differences of temporal patterns between islands and continental provinces	Timeline displays of events; interactive map-linked time series plots	Islands are disregarded as outliers
Differences in COVID-19 and Mobility behaviour along time (the three different waves)	Time series plots and cross-correlation analysis	Days since the start and duration of events added as features; separation of wave 1 from waves 2 and 3
Distinctive, non- generalisable COVID-19 and mobility patterns during wave 1.	Timeline view of misclassifications from the base model	The initial days of COVID-19 were cut off from the analysis

Table 2: Different types of knowledge generated using VA techniques and utilised in model development

patterns of the errors observed in M1 are not evident in M6. While each modification produced small improvements, the combination of all leads to significant performance enhancement. The accuracy improvement from about 0.71 to 0.86 was achieved through knowledge-informed operations modifying the input data for model training. The knowledge includes the initial understanding of possible interrelationships between population mobility and pandemic development, as well as specific insights about the processes in Spain during the COVID-19 pandemic. These insights were gained from the available data using VA techniques. Table 2 summarizes the types of knowledge extracted during data exploration and model building and how each piece of knowledge was utilized.

Our case study demonstrates the potential of integrating human knowledge into the data-driven model building process and the effectiveness of using visual analytics to gain relevant knowledge to guide this process.

4 General framework

In the previous section, we demonstrated how to incorporate human knowledge into modeling through a case study on COVID-19 and population mobility data from Spanish provinces. Building on this experience, we aim to generalize this model development process for application to similar temporal datasets.

4.1 Iterative development workflow

A strategy for advancing human-centered machine learning by integrating visual analytics with ML and explainable artificial intelligence (XAI) was presented in [2]. The key idea is that VA enables humans to understand and organize data and define relevant concepts, which can subsequently be used in ML algorithms. Human expert knowledge can then be utilized by the XAI component to provide meaningful explanations. In this paper, we focus on the initial part of this integration, developing a general workflow for incorporating human knowledge into an ML model.

Figure 8 illustrates a general workflow where the model is iteratively refined based on insights gained from the data or model results. In this framework, the knowledge acquired by humans is of primary importance. VA plays a dual role: enabling humans to build, extend, and refine their mental model of the phenomena and relationships in the data [1], and supporting communication between the human and the computer executing the modeling algorithm. While specifics may vary across different domains and applications, the types of the knowledge gained and the ways of using these knowledge types can be generalized, as discussed further in Section 4.2.

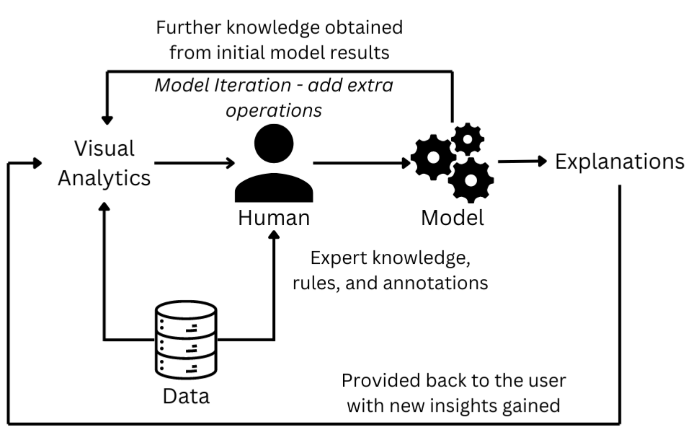


Fig. 8: Workflow for iterative model development involving human knowledge.

4.2 Knowledge-informed operations on input data

Human knowledge can be incorporated in a model through following generic operations on input data:

- **Define contexts:** For a model making predictions for time intervals, spatial locations, graph nodes, or other entities influenced by surrounding entities and/or conditions, identify the scope of the neighborhood where influences are substantial. Create a suitable representation of these contexts. For instance, in our case study, we defined temporal contexts of pandemic events, incorporating prior pandemic events and mobility behaviors.
- **Remove irrelevant patterns:** Eliminate patterns irrelevant to prediction tasks through smoothing, decomposition, or filtering. We removed the weekly variation cycle in the pandemic and mobility data using temporal smoothing.
- **Represent relevant patterns with descriptive features:** For time-variant attributes, represent the variation by summary features capturing value magnitude, development trend, and fluctuation intensity [3]. Similarly, other pattern types, such as dense spatial clusters or graph node communities, can be represented by summary features derived from the elements.
- **Temporal decomposition:** Apply when temporal variation patterns change over time due to seasonal trends, specific events, or interventions. In our study, we decomposed data by the 1st, 2nd, and 3rd waves to reflect different pandemic phases.
- **Spatial decomposition:** Applied to spatially distributed data to handle spatial disparities arising from natural or economic conditions, geographical barriers, or varying strengths of spatial links and communications.
- **Introduce features to capture trends:** Add features like 'days since the start' to help the model learn how phenomena change over time. For spatial trends, use distances from reference locations, such as transportation hubs or shopping centers. In graph data, similar features include distances from influential nodes, particularly, in social networks.
- **Remove outliers and unrepresentative data:** Exclude periods or anomalies that do not represent typical behavior. For example, in our study, we removed the initial days with patterns unique to this period.

By systematically applying these knowledge-informed operations, models can better capture underlying patterns and relationships in the data, enhancing performance across various applications.

4.3 Towards integration with XAI methods

Human-centered ML involves (1) integrating human knowledge into ML models and (2) explaining these models and their results meaningfully. Both aspects are essential for creating effective and trustworthy ML systems. This paper focuses on the first aspect. We describe how human expertise can be incorporated into ML models through feature engineering, data preparation and selection, and iterative model refinement.

For the second aspect, the same knowledge needs to be explicitly represented so that XAI methods can use it for generating understandable explanations. This can be a kind of meta-model with domain concepts and relationships. It may take the form of a knowledge graph [6] enhanced with human-oriented annotations. Extending VA systems to facilitate knowledge externalization and provenance capture is necessary for building such a meta-model [23].

However, a meta-model alone isn't enough for conducting explanatory conversations where users can ask for similar and contrasting examples and counterfactual explanations. For generating meaningful examples, the XAI system needs to understand data properties and constraints, such as temporal dependencies (e.g., COVID-19 morbidity levels cannot change instantaneously from c1 to c4). The system should also "know" what changes in the input are practically feasible for obtaining a desired outcome, like improving pandemic conditions by modifying mobility behavior. This kind of intelligence can be supplied to the XAI system as a combination of constraints.

Advancements in this direction require developing constraint typologies, constraint steerable XAI systems, and VA interfaces for capturing these constraints from experts during model building. These developments will enable ML systems to effectively integrate human knowledge and provide clear, actionable explanations, enhancing usability and trustworthiness.

4.4 Discussion

The approach taken in our case study mirrors real-world scenarios where expert knowledge and domain-specific insights are crucial for improving model performance. This aligns with many applications in fields such as healthcare, finance, and environmental science. We have demonstrated that human knowledge can be incorporated into a model even without developing specialized ML algorithms, but merely by modifying input data. Generalizing our experience, we created a set of widely applicable operations on data that can be used for integrating human knowledge into models irrespective of the ML methods employed for model training. This method-independence, combined with the existence of numerous tools in data science enabling the implementation of these operations, justifies their high practical value.

However, it should be acknowledged that the success of this framework depends on the availability and expertise of domain experts. Domain experts may not be skilled in conducting data science workflows, and the framework implies a collaborative approach where domain experts and data scientists work together, iterating on the model based on continuous feedback and insights. Ideally, this process should be supported by tools that are seamlessly integrated into the machine learning pipeline.

Another concern is that human knowledge can introduce biases. It is essential to validate and cross-check expert insights to mitigate the risk of biased outcomes. Tracking the provenance of all human-generated artifacts and recording the rationale for data modifications are also crucial for maintaining transparency and accountability.

Our endeavor to use the experience from a case study for developing a general framework appears both valid and feasible. The empirical success in improving model performance through human knowledge-driven operations provides a strong basis and motivation. By structuring this approach into a repeatable framework, we hope to offer a valuable methodology for many applications where expert knowledge is of primary importance.

5 Conclusions and Future Work

This paper presents a framework for integrating human knowledge into machine learning model development, demonstrated through a case study on COVID-19 and mobility data from Spanish provinces. By employing knowledge-informed operations on input data, we significantly improved model performance. Our approach is method-independent and leverages existing data science tools, making it broadly applicable across domains.

The key contributions of our work include: (1) a set of generalized operations for embedding expert knowledge into data preparation and feature engineering. (2) a practical workflow for human knowledge-driven model building; (3) empirical validation of the workflow and demonstration of its practical implementation.

Our results highlight the importance of human expertise in model refinement and the potential for further enhancing model interpretability and trustworthiness through the integration of explainable AI methods. The framework proposed in this paper targets key areas of human-centered AI, allowing human involvement in every stage of the modeling process. Furthermore, these methods have the potential to be extended to an XAI context where the obtained knowledge would form the basis for formulating explanations.

Our generalized framework for modeling has been proposed is currently supported by a single case study demonstrating its validity. The immediate future direction for this work is to test and refine the workflow based on other applications involving temporal phenomena, processes, or events. Additionally, this approach should be extended to include a workflow for providing explanations to end users.

References

1. Andrienko, N., Lammarsch, T., Andrienko, G., Fuchs, G., Keim, D., Miksch, S., Rind, A.: Viewing visual analytics as model building. Computer Graphics Forum **37**(6), 275–299 (2018). https://doi.org/10.1111/cgf.13324
2. Andrienko, N., Andrienko, G., Adilova, L., Wrobel, S.: Visual analytics for human-centered machine learning. IEEE Computer Graphics and Applications **42**(1), 123–133 (2022). https://doi.org/10.1109/MCG.2021.3130314
3. Andrienko, N., Andrienko, G., Artikis, A., Mantenoglou, P., Rinzivillo, S.: Human-in-the-loop: Visual analytics for building models recognising behavioural patterns in time series. IEEE Computer Graphics and Applications pp. 1–15 (2024). https://doi.org/10.1109/MCG.2024.3379851

4. Andrienko, N., Andrienko, G., Shirato, G.: Episodes and topics in multivariate temporal data. Computer Graphics Forum **42**(6), e14926 (2023). https://doi.org/10.1111/cgf.14926
5. Beckh, K., Müller, S., Jakobs, M., Toborek, V., Tan, H., Fischer, R., Welke, P., Houben, S., von Rueden, L.: Harnessing prior knowledge for explainable machine learning: An overview. In: 2023 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML). pp. 450–463 (2023). https://doi.org/10.1109/SaTML54575.2023.00038
6. Fensel, D., Simsek, U., Angele, K., Huaman, E., Karle, E., Panasiuk, O., Toma, I., Umbrich, J., Wahler, A.: Knowledge graphs. Springer Nature, Cham, Switzerland, 1 edn. (Feb 2020)
7. Giabbanelli, P.J., Jackson, P.J.: Using visual analytics to support the integration of expert knowledge in the design of medical models and simulations. Procedia Computer Science **51**, 755–764 (2015). https://doi.org/10.1016/j.procs.2015.05.195, international Conference On Computational Science, ICCS 2015
8. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**(3), 424–438 (1969), http://www.jstor.org/stable/1912791
9. Hong, H., Yoo, S., Jin, Y., Jang, Y.: How can we improve data quality for machine learning? a visual analytics system using data and process-driven strategies. In: 2023 IEEE 16th Pacific Visualization Symposium (PacificVis). pp. 112–121 (2023). https://doi.org/10.1109/PacificVis56936.2023.00020
10. Ponce-de Leon, M., del Valle, J., Fernandez, J.M., Bernardo, M., Cirillo, D., Sanchez-Valle, J., Smith, M., Capella-Gutierrez, S., Gullón, T., Valencia, A.: COVID-19 Flow-Maps an open geographic information system on COVID-19 and human mobility for Spain. Scientific Data **8**(1), 310 (Nov 2021). https://doi.org/10.1038/s41597-021-01093-5
11. Ponce-de Leon, M., del Valle, J., Fernández, J.M., Bernardo, M., Cirillo, D., Sanchez-Valle, J., Smith, M., Capella-Gutierrez, S., Gullón, T., Valencia, A.: COVID19 Flow-Maps daily cases reports (2021). https://doi.org/10.5281/zenodo.5217386
12. Ponce-de Leon, M., del Valle, J., Fernández, J.M., Bernardo, M., Cirillo, D., Sanchez-Valle, J., Smith, M., Capella-Gutierrez, S., Gullón, T., Valencia, A.: COVID19 Flow-Maps daily-mobility for Spain (2021). https://doi.org/10.5281/zenodo.5539411
13. Ponce-de Leon, M., del Valle, J., Fernández, J.M., Bernardo, M., Cirillo, D., Sanchez-Valle, J., Smith, M., Capella-Gutierrez, S., Gullón, T., Valencia, A.: COVID19 Flow-Maps population data (2021). https://doi.org/10.5281/zenodo.5226351
14. Liu, S., Andrienko, G., Wu, Y., Cao, N., Jiang, L., Shi, C., Wang, Y.S., Hong, S.: Steering data quality with visual analytics: The complexity challenge. Visual Informatics **2**(4), 191–197 (2018). https://doi.org/10.1016/j.visinf.2018.12.001
15. Liu, S., Wang, X., Liu, M., Zhu, J.: Towards better analysis of machine learning models: A visual analytics perspective. Visual Informatics **1**(1), 48–56 (2017)
16. Lu, Y., Garcia, R., Hansen, B., Gleicher, M., Maciejewski, R.: The state-of-the-art in predictive visual analytics. Computer Graphics Forum **36**(3), 539–562 (2017). https://doi.org/10.1111/cgf.13210
17. Lubba, C.H., Sethi, S.S., Knaute, P., Schultz, S.R., Fulcher, B.D., Jones, N.S.: catch22: Canonical time-series characteristics: Selected through highly comparative time-series analysis. Data Mining and Knowledge Discovery **33**(6), 1821–1852 (2019)

18. Muhlbacher, T., Piringer, H.: A partition-based framework for building and validating regression models. IEEE Transactions on Visualization & Computer Graphics **19**(12), 1962–1971 (dec 2013). https://doi.org/10.1109/TVCG.2013.125
19. von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Pfrommer, J., Pick, A., Ramamurthy, R., Walczak, M., Garcke, J., Bauckhage, C., Schuecker, J.: Informed machine learning – a taxonomy and survey of integrating prior knowledge into learning systems. IEEE Transactions on Knowledge and Data Engineering **35**(1), 614–633 (2023). https://doi.org/10.1109/TKDE.2021.3079836
20. Spinner, T., Schlegel, U., Schäfer, H., El-Assady, M.: explainer: A visual analytics framework for interactive and explainable machine learning. IEEE Transactions on Visualization and Computer Graphics **26**(1), 1064–1074 (2020). https://doi.org/10.1109/TVCG.2019.2934629
21. Wang, J., Liu, S., Zhang, W.: Visual analytics for machine learning: A data perspective survey. IEEE Transactions on Visualization and Computer Graphics pp. 1–20 (2024). https://doi.org/10.1109/TVCG.2024.3357065
22. Yuan, J., Chen, C., Yang, W., Liu, M., Xia, J., Liu, S.: A survey of visual analytics techniques for machine learning. Computational Visual Media **7**, 3–36 (2021)
23. Zhao, J., Glueck, M., Isenberg, P., Chevalier, F., Khan, A.: Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. IEEE Transactions on Visualization and Computer Graphics **24**(1), 340–350 (2018). https://doi.org/10.1109/TVCG.2017.2745279