J Vis
https://doi.org/10.1007/s12650-024-00989-9

REGULAR PAPER

Jiamin Zhu · Meixuan Wu · Yi Zhou · Nan Cao · Haotian Zhu · Min Zhu

Dowsing: a task-driven approach for multiple-view visualizations dynamic recommendation

Received: 21 November 2023 / Revised: 17 February 2024 / Accepted: 20 March 2024

Abstract Most users are able to obtain exploratory ideas from a data table but cannot clearly declare their analysis tasks as visual queries. Visualization recommendation methods can reduce the demand for data and design knowledge by extracting or referring information from existing high-quality views. However, most solutions cannot identify analysis tasks, which limits the accuracy of their recommendations. To address this limitation, we propose a deep learning and answer set programming-based approach to guide visualization recommendations by tracking potential analysis tasks and field preferences in exploration interactions. We demonstrate this approach via Dowsing, a mixed-initiative system for visual data exploration that automatically identifies and presents users' potential analysis tasks and recommends visualizations during exploration. Additionally, Dowsing allows users to confirm and edit their intentions in multiple ways to adapt to changing analysis requirements. The effectiveness and usability of our approach are validated through quantitative experiments and two user studies.

Keywords Visualization recommendation · Mixed-initiative · Dynamic task prediction

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12650-024-00989-9.

J. Zhu · M. Wu · Y. Zhou · H. Zhu · M. Zhu (✉)
College of Computer Science, Sichuan University, Yihuan Road, Chengdu 610065, Sichuan, China
E-mail: zhumin@scu.edu.cn

J. Zhu
E-mail: zhujiamin97@stu.scu.edu.cn

M. Wu
E-mail: wumeixuan@stu.scu.edu.cn

Y. Zhou
E-mail: zhouyi2@stu.scu.edu.cn

H. Zhu
E-mail: hao-tian.zhu@outlook.com

N. Cao
College of Design and Innovation, Tongji University, Fuxing Road, Shanghai 200092, China
E-mail: nan.cao@tongji.edu.cn

Published online: 17 April 2024

1 Introduction

Visual data exploration techniques aim to leverage human's powerful visual capability to help users generate hypotheses about data and enhance the efficiency of data analysis (Ferreira de Oliveira and Levkowitz 2003; Shirato et al. 2023). Visualization recommendation techniques are methods that take user-uploaded datasets and optional configuration parameters as input to automatically generate high-quality charts, which can reduce the demand for visualization design knowledge and improve the efficiency of data exploration (Zhu et al. 2020). However, during data exploration, users often hold vague analysis tasks that may dynamically change as the exploration progresses (Gotz and Zhou 2008; Battle and Heer 2019). The effectiveness of visualization varies depending on tasks (Kim and Heer 2018; Saket et al. 2019; Sarikaya and Gleicher 2018), and existing visualization recommendation systems often struggle to dynamically identify users' analytical tasks, thereby limiting the quality of their visualization recommendations. Meanwhile, users may iteratively construct multiple visualizations to fulfill complex analytical tasks during data exploration(Yu et al. 2023; Roberts 2007). Most visualization recommendation systems fail to capture the connections among the creation processes of multiple views (Wu et al. 2021), further limiting their usability. This motivates us to propose an approach that can leverage the actions of users creating and editing multiple views throughout the visual data exploration process, dynamically identify users' analysis tasks during visual data exploration and generate targeted visualization recommendations based on the identified tasks.

Some studies on visualization recommendation employ task-based approaches to either constrain or enhance the recommendation outcomes. Draco Moritz et al. (2019) specifies two task categories, namely value and summary, that users can declare in the configuration to obtain more appropriate visualization recommendations. NL4DV (Narechania et al. 2021) extracts analytical tasks from natural language inputs and sets recommendation strategies based on the identified tasks. However, these methods require users to specify analytical tasks, and the ambiguity of analytical tasks in data exploration often poses a challenge for users to classify their aims into predefined categories or formulate normative natural language queries. Alternatively, other studies such as BDVR Gotz and Wen (2009), VisGuide Cao et al. (2022), and View-seeker Zhang et al. (2021) leverage provenance data to analyze user preferences and behavioral patterns to enhance the effectiveness of visualization recommendation. Nevertheless, significant differences exist between "user preferences or behavioral patterns" and "analytical tasks," which hinders the full utilization of existing research on the effectiveness of task and visual encoding in the visualization community, thereby limiting the effectiveness of their visualization recommendations.

In order to further improve the effectiveness of visualization recommendation methods for data exploration, this paper proposes to extract users' analytical tasks and field preferences from action logs using deep learning-based techniques. Subsequently, visualization charts will be recommended based on the extracted analytical tasks and field preferences. The undertaking of the aforementioned tasks faces two primary challenges:

Identifying uncertain and dynamic analytical tasks from provenance data presents the first challenge, as the action log often contains a substantial amount of irrelevant noise(Xu et al. 2015). Task identification algorithms need to be robust to noise and ensure the reliability of their results. Moreover, users' analytical tasks tend to be stable over a short period of time but can change gradually during exploration. Therefore, task identification algorithms should be able to adapt to such changes.

Generating visualization charts based on analytical tasks and field preferences poses the second challenge. Despite lots of studies on the relationship between tasks and the effectiveness of visual channels (Kim and Heer 2018; Saket et al. 2019), task-specific visualization design principles are relatively scarce and not easily applicable to constructing independent visualization recommendation methods. Moreover, while previous studies have qualitatively explored the correlation between task requirements and the effectiveness of visual channels, further investigation is needed to determine how these findings can be integrated for the purpose of visualization recommendations.

To address the above challenges, we propose a task-centered visualization recommendation workflow. The workflow consists of two stages: a human–computer mixed agency and a visualization recommender. We begin by summarizing common analytical tasks in data exploration through literature review and discussions with analysts. To overcome the scarcity of appropriate training data, we conduct a semi-structured workshop to obtain a corpus of "action logs-task" pairs. For stage 1: **A human–computer mixed agency is designed to address the first challenge**. We utilize a deep learning model based on long short-term memory (LSTM) Hochreiter and Schmidhuber (1997) to classify users' action logs into multiple analytical tasks and dynamically identify the latest actions using a sliding window strategy. Simultaneously, the human–computer mixed agency employs a Chi-square-based method to detect users' field preferences. We also mix the raw output of the human–computer mixed agency with its historical output to enhance short-term temporal stability, and take into account the impact of multiple views on the analytical task. For stage 2: **A visualization recommender is implemented to address the second challenge**. It is implemented using answer set programming, which encodes universal visualization design rules based on Draco Moritz et al. (2019) and task-specific visualization design rules from existing research. We design a ranking and deduplication strategy to remove visually similar charts from the recommendation results and maintain conciseness. Finally, we propose Dowsing, a mixed-initiative visualization recommendation system, to support analysts in conducting visual data exploration. In summary, our contributions are as follows:

- We identify common task categories in data exploration through expert interviews and, for the first time, collect a real-world "action log-task" dataset via a workshop.
- We propose a method that combines a deep learning-based model and statistical testing to dynamically identify analytical tasks and field preferences from users' action logs. Through the design of input features in our model, we have further enhanced the robustness of our method.
- Building upon existing research, we propose a task-centered visualization recommendation approach to recommend visualizations that meet specified tasks. Additionally, we provide a mixed-initiative recommendation system for visual data exploration and validate its usability through quantitative experiments and user studies.

Dowsing has been open sourced at GitHub (https://github.com/Dowsing-Machine/Dowsing), and a live demo can be found at http://dowsing-machine.com.

2 Related work

The related work of this paper involves three areas: visual analysis tasks, visualization recommendations, and interactive analysis in data exploration.

2.1 Visual analysis tasks

Early works on automatically generating information presentations (Sarikaya and Gleicher 2018; Roth and Mattis 1990; Collins et al. 2018) categorized and described user intent in a way that could form a preliminary summary of analytical tasks. Further studies delved into analytical tasks, offering more formal definitions and classifications. Amar et al. (2005) proposed ten low-level analytical activities as components that have been widely adopted in subsequent research. Rind et al. (2016) categorized user tasks based on abstract, composite, and perspective dimensions and suggested using "objective" and "action" to describe user tasks. Brehmer and Munzner (2013) proposed a hierarchical visualization task classification system that can represent complex tasks as sequences of interdependent components.

Furthermore, research has established a correlation between analytical tasks and visual design. User studies conducted in some studies (Kim and Heer 2018; Saket et al. 2019; Luo et al. 2019) have shown that chart types and visual encodings can influence the effectiveness of tasks. Quadri and Rosen (2022) arrived at similar conclusions by summarizing existing research and offered suggestions for visualization design specific to tasks. Albers et al. (2014) examined how visual encodings affect various aggregation tasks. TaskVis Shen et al. (2021) summarized research on the relationship between tasks and visual design and developed a set of rules to represent them.

Considering the crucial contribution of analytical tasks in guiding visualization design during data exploration, we identify common analytical tasks through expert interviews and integrate them into the visualization design process by leveraging previous research.

2.2 Visualization recommendation

Visualization recommendation systems, which are primarily designed for data exploration purposes, usually suggest visualizations based on provided datasets and optional user specifications (Soni et al. 2024). DeepEye Luo et al. (2018) proposed three criteria to assess the quality of a visualization and employed a partial-order-based method to identify effective candidate visualizations.

Voyager Wongsuphasawat et al. (2017) utilized the CompassQL (Wongsuphasawat et al. 2016) recommendation engine, which employs data attributes and perceptual principles to enumerate, cluster, and rank visualizations. Draco Moritz et al. (2019) encoded visualization design principles as rules and employed answer set programming to generate visualizations. KG4Vis (Li et al. 2022) applied embedding techniques based on Trans-E to learn entity and relation embeddings from dataset-visualization pairs and leveraged a knowledge graph to provide visualization recommendations. Zhu et al. (2023) propose a method based on a mathematical programming model that can automatically generate timeline layouts with reduced visual clutter, thus assisting users in identifying hidden patterns. Zhou et al. (2023) developed an approach that employs saliency maps to automatically produce text descriptions from visual images, facilitating the process of data analysis. Additional literature (Zhu et al. 2020; Qin et al. 2020; Wu et al. 2022) offers a more extensive overview of visualization recommendation systems.

With the increasing refinement of user demands, there has been a growing trend of integrating analytical tasks into visualization recommendation systems. TaskVis (Shen et al. 2021) provided a summary of low-level analytical tasks and their corresponding chart types, and established an analytical task model to implement task-oriented recommendation engines. NL4DV (Narechania et al. 2021) integrated five low-level analytical tasks into a Python package that infers tasks from natural language queries and created a visualization toolkit. Furthermore, Shi et al. (2019) proposed a reinforcement learning-based sequential chart generation method and trained optimization models for three major analytical tasks.

In contrast to existing visualization recommendation methods that lack support for allowing users to specify analysis tasks or necessitate users to explicitly specify analysis tasks, our proposed approach provides a comprehensive consideration of users' analytical tasks, and recommends visualizations to satisfy requirements based on task types and field preferences.

2.3 Interaction analysis in data exploration

During data exploration, users generate a substantial number of interactions. Provenance data researchers suggest that these interactions may contain valuable information concerning users' personalities, task execution efficiency, and insight discovery (Gotz and Zhou 2008; Quadri and Rosen 2022; Guo et al. 2016; Dabek and Caban 2017; Nguyen et al. 2020; Rubab et al. 2021). Stoiber et al. (2022) proposed an interaction assistance conceptual model for Visual Analytics systems to help domain experts effectively use and understand visual representations through onboarding and guidance interactions. Gotz and Wen (2009) presented four common patterns of analytical behavior in regular expression form and proposed a visualization recommendation method based on these patterns. Brown et al. (2014) utilized machine learning techniques to predict users' task execution speed and personality factors based on their interaction data. Li et al. (2022) applied LSTM to model users' interaction behavior when using multi-view visualizations and provided interaction recommendations for novice users. VisGNN Ojo et al. (2022) extracted embeddings of users, visualization configurations, and attributions from historical data to predict specific users' ratings of visualization designs. VisGuide Cao et al. (2022) collected user preferences on granularity, consistency of generation operations, encoding transformations, and other aspects through implicit labeling mechanisms and trained an online linear regression model for visualization recommendation.

Overall, current research on interactive analysis to support data exploration primarily focuses on analyzing users' preferences, Qian et al. (2022) considered data preferences and visual preferences as key factors in personalized visualization recommendation. However, users often lack professional skills to translate their intentions into visual configurations, which results in visual preferences obtained from user interactions may not always be accurate. As another critical aspect of users' preferences, data preferences, refined into field preferences, are taken into consideration in this paper. Meanwhile, we introduce analytical tasks as a joint outcome of interactive analysis to enhance the interactivity of the recommendation results.

3 Dowsing design

In this section, we discuss the design goals of Dowsing and the analytical tasks supported by Dowsing. They are determined through expert interviews and literature research. Additionally, a usage scenario is presented to demonstrate the practical application of Dowsing.

3.1 Design goals

In order to improve the design of Dowsing's visualization recommendation workflow, we conducted interviews with three data analysis experts (referred to as E1, E2, and E3) to gain insights into their requirements during visual data exploration. We obtained their requirements and suggestions for visualization recommender systems during the data exploration process. All experts have more than three years of experience in data analysis and are not affiliated with the authors of this paper. Combined with our analysis of existing visualization recommendation and data exploration systems, we have identified the following key design goals:

G1. Consider both data characteristics and user tasks simultaneously. Users may have vague analytical tasks during data exploration (Gotz and Zhou 2008), which should be utilized to guide visualization recommendations. Particularly, E2 pointed out that the distribution of data may not always align with the user's expectations, leading to potentially inaccurate analytical tasks. Therefore, generating visualization recommendations requires considering the data characteristics and the user's tasks simultaneously.

G2. Adapt to users' evolving analytical requirements while maintaining stability. Although users' analytical requirements may remain stable over a short period of time during data exploration, these requirements may change as the exploration progresses. E1 noted that they do not frequently alter their assumptions about the data or analytical intentions, but may do so if the results are interesting. Therefore, Dowsing's identification of analytical tasks should remain consistent over time but be flexible enough to adapt to changes in exploratory behavior. Additionally, Dowsing's visualization recommender should be able to dynamically respond to these changes in real-time.

G3. Detecting field preferences of users from the data to refine the recommendations. The analysis of a dataset often involves multiple data fields, but users may only be interested in a subset of them. E1 suggested that analysts typically select a few fields for exploratory analysis after reviewing the data. E3 mentioned that although recommending fields favored by analysts can help them focus on the current analytical task, it may also limit their exploration of other data fields. Therefore, it is important for Dowsing to carefully identify user preferences for data fields and use them to refine the recommendation results as appropriate.

G4. Support user comprehension and interactive control of the recommendation process. Algorithms and models may not consistently generate optimal results, necessitating human supervision to establish a closed-loop control and ensure the recommendation workflow stays on course. E3 highlighted the significance of control in data exploration systems, which should enable analysts to iteratively fine-tune the system to align with their requirements. Hence, Dowsing ought to uncover the crucial parameters governing recommendation outcomes, and enable users to effortlessly customize them to satisfy their needs.

G5. Facilitates the organization and arrangement of charts. During the process of visual data exploration, analysts frequently generate multiple charts to carry out distinct analytical tasks. However, experts indicate that the exploration process is not always structured, implying that charts may be created in a random sequence, with consecutively produced charts potentially addressing different analytical tasks. E1 and E2 mentioned that they often need to collate charts that serve similar analytical purposes into a designated area, positioning them spatially adjacent to one another to facilitate cross-comparison of various views, thereby enhancing the efficacy of the analytical tasks.

3.2 Modeling analytical tasks

In order to enhance the usability of Dowsing, we collaborate with experts by conducting interviews to identify the analytical tasks that Dowsing supports. We initially derive our task list from the 18 analytical tasks summarized by Shen et al. (2021). We exclude the Spatial, Part to Whole, and Error Range tasks, which typically necessitate specialized data types and display formats. Combining expert opinions, we consolidate similar tasks and classify infrequent tasks as "others". Additional information is available in the supplementary materials. Ultimately, we compile a task list consisting of six tasks, namely Confirm, Transform, Trend, Compare, Correlation, and Others. Their descriptions are summarized in Table 1.

Table 1 The 6 analytical tasks supported by Dowsing, and their examples

Task	Example
Confirm	What's the precipitation in New York on July 1st?
Transform	What's the highest temperature per year?
Compare	Is there a significant difference in temperature New York and Chicago?
Trend	How does the average monthly precipitation change over the year?
Correlation	Is there any relationship between wind speed and precipitation?
Others	Is there any extreme high temperature in the data?

3.3 Usage scenario

We illustrate how Dowsing enhances the efficiency of visual data exploration via task-driven visualization recommendations, using a concise usage scenario. This scenario is also showcased in the supplementary video (available online at https://dowsing-machine.github.io) and can be entirely accomplished using Dowsing.

Theodore is a data analyst who possesses fundamental knowledge of visualization and intends to analyze the weather data of two cities and extract valuable insights from it. In order to comprehend how Dowsing can assist users in data exploration, we present his workflow in a sequential manner.

Theodore initiates his analysis by opening the dataset using Dowsing and examining the data fields. He notes the presence of the 'Year', 'Month', and 'Day' fields and considers the possibility of any underlying **trends** within the data. To investigate the presence of significant trends, Theodore creates a chart (Chart#0) in which he opts for a scatter plot as it can effectively display individual data points. He chooses 'Year' as the X-axis and experimentally selects 'precipitation' as the encoding field for the Y-axis. Additionally, Theodore uses 'location' as the color encoding to differentiate the data from the two cities and applies the 'mean' aggregation function to 'precipitation'. Although the scatter plot distinguishes the difference in annual average temperature between the two cities, it does not effectively show the time trend. As a result, Theodore switches to a line chart to display the temporal variation of annual average precipitation.

Dowsing's task prediction module suggests two potential analytical tasks, **comparison** and **trend**, by analyzing historical actions. Theodore concurs with the prediction and seeks chart recommendations for similar insights. He observes two intriguing recommended results, Chart#1 and Chart#2, which, respectively, display the trend of minimum temperature and precipitation by monthly average for the cities. In comparison with Chart#0, they more effectively illustrate the differences in weather trends between the cities. Therefore, Theodore replaces Chart#0 with these visualizations.

Besides comparing the differences between cities, Theodore also desires to investigate the potential correlations among weather indicators. He constructs a scatter plot (Chart#3), utilizing precipitation as the X-axis encoding and wind as the Y-axis encoding. Theodore encounters difficulty in obtaining insights from the current visualization results. Based on Theodore's latest historical actions, Dowsing infers that he may be involved in **comparison** and **transformation** tasks. However, since Theodore believes that the predicted tasks may not fully align with his requirements, he retains the comparison task and substitutes the data transformation task with **correlation**. By browsing the recommended results, Theodore uncovers that utilizing color encoding for the weather field in the visualization (Chart#4) reveals that both precipitation and wind are associated with weather, with rain frequently accompanied by strong winds.

Theodore desires to delve deeper into the temporal and spatial attributes of the high precipitation data points in Chart#4; therefore, he generates Chart#5, which employs a scatter plot to visually encode the relationship between the highest temperature of cities and the Month. By leveraging cross-view brushing, he identifies that these data points are all sourced from New York in the spring and summer seasons.

After conducting additional data exploration, Dowsing detects Theodore's preferences for Year, precipitation, and location and initiates field preference detection to offer more personalized recommendations. Theodore peruses the recommended results and refines the predicted tasks to align with his requirements, leading to a satisfactory visualization (Chart#9) that exhibits the yearly fluctuations in precipitation.

4 Workshop

Machine learning algorithms require training data for support. Currently, there is no dataset available to support task-driven visualization recommendations. To address this issue, we have organized a workshop to collect the necessary data for Dowsing's task prediction. In this section, we present the design of the workshop and report on our preliminary analysis of the collected data.

4.1 Procedure

We developed a basic data exploration system beforehand to capture user actions during the workshop. This system resembles Dowsing's user interface, but does not provide task-driven visualization recommendation. User actions during the data exploration process were collected through event tracking. Further information regarding this system can be found in the supplementary materials. To conduct the experiment, we selected two real-world datasets from vega-dataset that with moderate data sizes and have both quantitative and categorical data types.

We recruited sixteen university students, comprising ten males and six females, with diverse experiences in visualization and data analysis to participate in the workshop. None of them had prior knowledge of the datasets.

During the workshop, we conducted a basic tutorial (30 min) to introduce the concepts of visualization, the definitions of involved analytical tasks and the usage of our experimental system through cases. Subsequently, we allocated 10 min to address any inquiries from the participants regarding the experimental system and tasks. Finally, we instructed the participants to explore two datasets, create visualizations for each dataset, and stop when they believed that they had completed the data exploration. Participants were required to annotate the reasons for each visualization creation and the corresponding completed analytical tasks in natural language. The exploration for each dataset lasted approximately 40 min.

4.2 Result

During the workshop, the participants' actions were recorded, which encompassed the modification of **visual properties** (chart types and visual encoding), **data processing** (aggregation, filtering, and binning), and **system control operations** (annotation, switching datasets, and undo-redo). A total of 194 valid charts and 2,658 actions were collected. Task labels were manually assigned to each chart based on participants' annotations, yielding 450 task labels across all 194 charts.

In the workshop, we observed that individual participants can numerous noisy actions while exploring data by habit, that can disrupt the model's task prediction performance. Therefore, we decide to identify action types that are significantly correlated with tasks by Chi-square tests and use them to construct the model's feature representation that enhancing robustness. Details of the statistical results are available in Appendix 3. Each chart is decomposed orthogonally into twelve configurations for visualization and data processing, where four configurations were revealed significantly associated with the task: **chart types** (bar, line, and scatter), **x-axis aggregation**, **y-axis aggregation**, and **x-axis data type**. These findings suggest that actions associated with these chart configurations may have a greater impact on tasks.

Subsequently, we partition the action log into 194 action sequences based on the corresponding charts. These action sequences share the same task labels as their respective charts and serve as the input to train the proposed task prediction model.

5 Dowsing

We propose Dowsing, a mixed-initiative recommendation system for visual data exploration that dynamically predicts and displays users' analytical tasks during their exploration process, and generates recommendations based on these tasks. As shown in Fig. 1, Dowsing comprises two stages: a human–computer mixed agency and a visualization recommender. The human–computer mixed agency is utilized to identify users' analytical tasks and field preferences dynamically. It is additionally designed to enhance task prediction stability and enable manual adjustments of the agency's output. The visualization recommender applies answer set programming, which synthesizes the statistical characteristics of the dataset and the agency's output, namely the analytical task and field preference, to generate visualization recommendations. Dowsing's recommendation workflow is extensible, allowing researchers to support more analytical tasks or improve recommendations through updated research without altering Dowsing's overall design.

In this section, we present the human–computer mixed agency and visualization recommender of Dowsing, followed by a description of its user interface.

5.1 Human–computer mixed agency

The human–computer mixed agency generates facts through three independent components: data fact extraction, task prediction, and field preference detection.

5.1.1 Data fact extraction

To satisfy G1, we employ data fact extraction to obtain metadata and statistical features from the given dataset, providing essential information for the recommender. We expand upon the data fact representation in Draco Moritz et al. (2019), and measure data features including data type, entropy, mean, median, and range for each field. Additionally, we calculate Pearson correlation coefficients between all numerical fields, which can be applied to establish rules for correlation tasks.

5.1.2 Task prediction

Task prediction involves predicting the analytical tasks being performed by users based on their action sequences, which is a supervised sequence-multi-label classification problem since users may engage in multiple tasks concurrently.

As discussed in 4.2, it is crucial to select relevant features from the action log for robustness. Drawing on the data collected from workshop, we select 30 distinct user actions, including visual channel and chart type modifications, data aggregation changes, etc. (see supplement for full list). Each action is represented as a one-hot encoding. For the sake of G2, we incorporate the latest actions into feature $\mathrm{F}_{\mathrm{raw}}$ via sliding windows, where a window size of 15 covers the actions (13.7 in average by workshop) that is needed to create a visualization. Such a window size enables the model to capture information on slightly more than one chart creation (Fig. 2).

We also devise features highly correlated to tasks, aiding the model against noise. Based on the relevance analysis in Sect. 4.2, chart type, aggregation, and channel data types are crucial features for tasks. Thus, we design another feature $\mathrm{F}_{\mathrm{grouped}}$, also sized 15 like $\mathrm{F}_{\mathrm{raw}}$. $\mathrm{F}_{\mathrm{grouped}}$ first categorizes actions into four types: chart type (T), aggregation method (A), encoding channel data type (E), and others (O). The size of each group is set to LT=7, LA=2, LE=2 and LO=4, following the proportion of actions from the workshop. Each group retains the corresponding most recent actions. When the action sequences are not enough fill into $\mathrm{F}_{\mathrm{raw}}$ and $\mathrm{F}_{\mathrm{grouped}}$, we inset Empty actions in remaining slots. Eventually, $\mathrm{F}_{\mathrm{raw}}$ and $\mathrm{F}_{\mathrm{grouped}}$ are concatenated into the final input feature.

Taking into account of limitations of the volume of training data and the computational resources, we choose long short-term memory (LSTM) as the backbone network over Transformer. Bidirectional LSTM (Bi-LSTM) Schuster and Paliwal (1997) network enhances bidirectional context modeling over regular LSTM, is more appropriate for modeling users' contextual actions. Therefore, we adopt it to learn the mapping from sequences to tasks. As illustrated in Fig. 3 and Eq. 1, the constructed Bi-LSTM-based model outputs concatenates forward and backward hidden states, passing through a flattening layer and sigmoid-activated linear layer to derive predictions, and is trained via cross-entropy loss and the Adam optimizer.

$$\widetilde{T} = \text{Bi-LSTM}(F_{raw}||F_{grouped}) \tag{1}$$

The output $\widetilde{T}$ is a 6-dimensional vector, with each dimension corresponding to the confidence level for a task. For example, a higher value of $\widetilde{T_0}$ signifies a greater likelihood of the existence of Task 0. Furthermore, to improve stability and incorporate historical information (G2), the prediction output is smoothed with the last one, which can be formatted as

$$T = \alpha * \widetilde{T} + (1 - \alpha) * T_{last} \tag{2}$$

where T denotes the final output of the task prediction component. To address G4, users can override T's default value through the interface to specify requirements. For instance, users can lock a task in an activated state, fixing its corresponding T dimension at 1; or mask a task, forcing its T dimension permanently to 0.

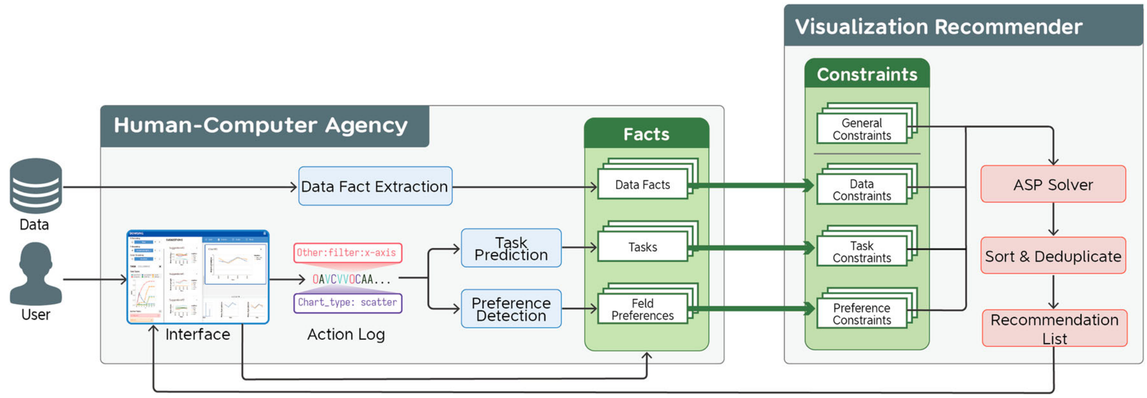

Fig. 1 Dowsing's workflow. The workflow consists of two main stages: (1) The human–computer mixed agency generates data facts, task facts, and field preferences while accepting user control. (2) The visualization recommender subsequently ingests these facts to produce visualization recommendations via answer set programming (ASP) rules

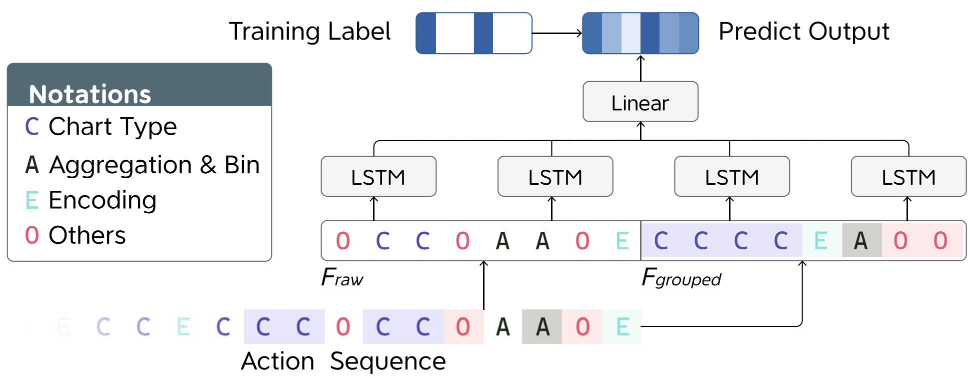

Fig. 2 Dowsing's task prediction model. The action logs are encoded by one-hot representation and are processed simultaneously into two input features, F_{raw} and $F_{grouped}$, for the Bi-LSTM model. The output is a 6-dimensional vector denoting the likelihood of the existence of each task

5.1.3 Field preference detection

The field preference detection analyzes action sequences to construct access a frequency table for the data fields. Field access frequencies represent user preferences, with more frequent fields having greater weights in the recommender and higher chances of appearing in recommendations. For G3, Chi-square tests are applied to activate preference detection only when a distinct field preference exists. However, users can manually override this setting.

Actions adding or assigning data fields to visual channels are defined as 'accesses' to those fields. For instance, when a user modifies a chart by changing the 'x-axis' encoding from 'Field A' to 'Field B', it constitutes an access to 'Field B'. Since individual field accesses can represent samples of user's field preference, Chi-square tests can determine if users have significant preferences for particular fields.

Following each field access, we first confirm a minimum sample of 40, then fit the users' field access frequencies against equal probability accesses using a Chi-square test. If the p-value is under 0.05, the user is considered to have significant preferences for certain fields. In such cases, a rule is generated encouraging recommended views to contain above-average accessed fields. As with the task prediction component, users can actively adjust field preferences through the interface (G4).

5.2 Visualization recommender

The recommender comprises two components: a visualization solver utilizing answer set programming (ASP)-based visualization rules, and strategies for ranking and deduplication. In this section, we first briefly introduce answer set programming and then explain how these rules are applied in the recommender. The ranking and deduplication component implemented in Dowsing are also described.

5.2.1 Answer set programming

Answer set programming Vladimir (2008) is a declarative programming paradigm first proposed by Dimopoulos et al. for solving complex search problems. In ASP, search issues are encoded as programs executed by answer set solvers. ASP programs comprise atoms, literals, and rules. Atoms are basic propositions that can be true or false. Literals are atoms or their negation. Rules contain literals, e.g., `A:- L1, L2, L3` derives A as true when `L1, L2, L3` are all true; otherwise, `A` is false (violates the rule). An atom alone can represent an unconditional fact. Rules that must be satisfied are hard constraints. Soft constraints are optional but have weights denoting violation costs. The cost of a result is the sum of weights of violated soft constraints, representing the quality of result.

5.2.2 Rule for visualization recommendation

ASP is used to represent visualization rules due to (1) the extensibility, which allows reusing or adding rules without fully rebuilding the recommendation algorithm; and (2) existing research Moritz et al. (2019) has shown ASP has superior computational performance compared to heuristic algorithms when handling numerous rules.

Dowsing's visualization rules comprises three parts: general rules ensuring basic visualization quality; task-specific rules measuring support for analytical tasks; and preference rules elevating user-preferred fields in recommendations:

- The visualization recommendations are derived from Draco Moritz et al. (2019)'s ruleset but impose additional constraints on the design space regarding supportable chart types, visual channels, and data transformations. These rules serve as the fundamental and fallback recommendation rules of the system. They are capable of providing users with a consistent visualization recommendation experience, even in situations where task prediction information is absent (a scenario that typically occurs in the early stages of visual data exploration, also known as the cold-start phase). These early recommendations help users to quickly begin exploring data and become familiar with the use of Dowsing.
- Expanding on visualization principles for analytical tasks compiled from TaskVis (Shen et al. 2021) and related works (Kim and Heer 2018; Saket et al. 2019), we formulated 31 additional rules for tasks. For instance, to aid the 'determine value' task, we imposed a soft constraint promoting bar charts.
- Preference rules act as soft constraints, reducing the loss for visualizations encoding the user's preferred fields. We use soft constraints instead of hard ones to measure the differences in how often two fields are accessed. Soft constraints also allow for visualizations that don't strictly follow the rules, which means we can recommend options that include fields the user hasn't looked at yet, making the suggestions more varied.

5.2.3 Sorting and deduplication

After the ASP solver resolves the answer set according to the rules, we obtain all the "visualization candidates". These candidates are then sorted in ascending order of the cost summarized by ASP, and the 20 candidates with the lowest cost are selected. To ensure conciseness of the results, the candidates are grouped by chart type and visual channel, and visualizations within the same group are considered similar. For similar visualizations, only the one with the minimum cost is retained.

5.3 Interface

Here, we present the interface and functionalities of Dowsing. Dowsing is a mixed-initiative system that continuously analyzes user tasks to enhance the efficiency of visual data exploration. Its non-intrusive design allows users to freely choose whether to explore and adopt recommended results. We highlight that in Dowsing, users have the flexibility to adjust the tasks and field preferences predicted by the agency for better satisfy their needs (G4).

Figure 3 shows the interface of Dowsing, which includes a control panel (Fig. 3A–C) for displaying and controlling charts and agency. The suggestion panel (Fig. 3D) lists the visualization recommendations. The canvas (Fig. 3E) allows users to drag and customize the layout of charts to support G5. The history panel (Fig. 3F) exhibits the editing timeline for the chosen chart. Raw data columns appear as in Fig. 3G, where users can also switch datasets or upload their own datasets.

5.3.1 Control panel

The control panel can be further divided into two parts: the view controller and the agency controller. Specifically, view controller A enables users to adjust the chart type and visual encoding (x, y, and color) of the selected visualization, as well as perform aggregation, filtering, and binning operations on the data fields associated with each visual channel.

The agency controller enables viewing and modifying the predicted tasks (B) and field preferences (C) to align with user requirements. A line chart depicts the real-time impact of actions on task scores, with active task tags below adjusting and reordering concurrently (G2). Users can disable tasks temporarily by clicking tags to compare recommended visualizations across tasks. Permanent task removal or assignment is also allowed. Likewise, the preference control displays a dynamic bar chart of field access frequencies. Preferred fields appear above ordinary fields in distinct colors. Field labels allow manually pinning/unpinning preferences. Further, the agency settings can be used to switch on and off the task prediction, preference detection, and recommender.

5.3.2 Suggestion panel

The suggestion panel exhibits recommender-generated visualizations. Recommendations can be inserted into the canvas by clicking on the Add button for each candidate. Users can collapse the panel when unnecessary to obtain more canvas area.

5.3.3 Canvas

The canvas uses a 12-column grid system for creating and editing multi-view visualizations, which are appeared as draggable, resizable, or removable grid elements. We choose a grid layout since most multi-view visualizations align to a finite grid Chen et al. (2021). Grids minimize unnecessary alignment and movement while enabling extensive customization. Visualizations adapt layouts to viewport changes. Brushing propagates to charts sharing fields, enabling cross-view highlighting. Users can manually annotate visualizations for personal analysis. With permission, annotations would be used for further improvement of the task prediction model.

5.3.4 History panel

The history panel aids users in reviewing the editing process of a visualization and permits it to be reverted to any prior state. The historical states of the visualization are depicted by thumbnails. Similar to the suggestion panel, it can be collapsed to expand the canvas area.

6 Evaluation

We begin by assessing the effectiveness of the task prediction model through quantitative comparisons and ablation study. Subsequently, we explore the effectiveness of Dowsing through user studies.

Par95 https://hmmlearn.readthedocs.io/en/latest/.

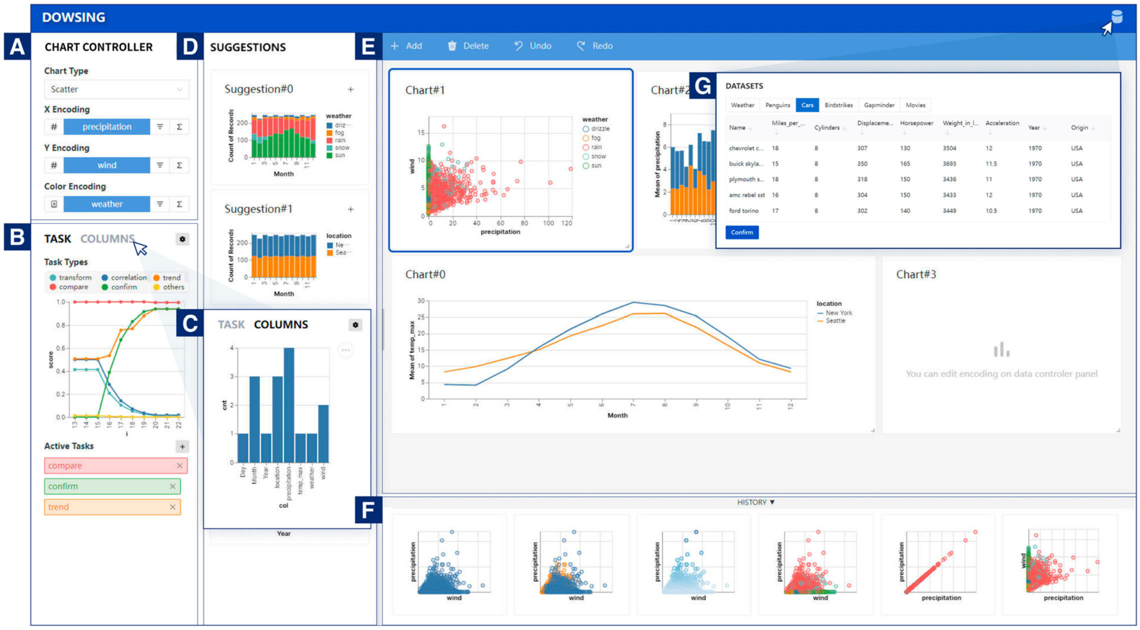

Fig. 3 Dowsing's user interface. The chart controller **A** permits manipulating visual encodings and types of a selected chart. The agency controller **B/C** allows observing and controlling the identified tasks and field preferences. The suggestion panel **D** displays recommender-generated visualizations in ascending cost order. The canvas **E** enables creating, editing, and organizing charts. The history panel **F** permits reviewing charts' edit timelines. The raw data table appears in **G**

6.1 Evaluation of task prediction model

We quantitatively assess our task prediction model's accuracy and evaluate the impact of feature designs on robustness via ablation study. The dataset is collected during the workshop and contains a total of 2658 "action log-task" pairs.

Quantitative analysis and comparison. We utilize the average accuracy of 10 runs of 5-fold cross-validation on the workshop dataset as the evaluation metric and compare it with five relevant machine learning baselines. They are hidden Markov model (HMM) [1] Rabiner and Juang (1986), decision tree (DT) [2] Song and Ying (2015), k-nearest neighbors (KNN) [3] Peterson (2009), random forest (RF) [4] Breiman (2001), and multilayer perceptron (MLP) [5] Popescu et al. (2009) classifiers. The findings indicate that our Bi-LSTM-based model outperforms other machine learning models in terms of accuracy by at least 2.2%.

Feature ablation study. To verify the effectiveness of our feature selection strategy, we conduct ablation studies on F_{raw} and $F_{grouped}$. In addition, the model's robustness to noise should be considered. We randomly selected some actions (5% or 10% of the total) during testing and added extra noise by replacing the same number of other actions, to compare the robustness of these models. The results validated the effectiveness of our feature selection strategy and the model's noise resistance ability. Figure 4 shows the performance improvement brought by the concatenation of F_{raw} and $F_{grouped}$. Although the extra noise led to drops in accuracy, the performance remained acceptable.

6.2 User study

We performed two user studies assessing Dowsing's effectiveness and usability. In the first study, a structured experiment was performed to compare Dowsing against alternative visualization recommenders

Par95 https://hmmlearn.readthedocs.io/en/latest/.

Par96 https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html.

Par97 https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html.

Par98 https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.

Par99 https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html.

and their impacts on data exploration efficiency. In the second study, we invited users to explore data with Dowsing, analyze their usage patterns, and collect subjective feedback.

6.2.1 User study 1: comparison with alternative systems

We first performed a structured user study to address: "*Relative to alternative systems, can Dowsing enhance data exploration efficiency?*". This question is critical, as it validates the core recommendation approach of Dowsing. We compared it against Voyager2 as an alternative system, which supports data exploration similarly and provides targeted recommendations based on the current view.

Participants. We recruited twelve participants from university, consisting of three graduate students and nine undergraduate students (eight males, four females) who possessing experience in data analysis. We confirmed that participants had never been exposed to the dataset or system used in the study.

Datasets. We chose two real-world tabular datasets, Iris and Driving, each consisting of five and four fields correspondingly.

Procedure. We first introduced Dowsing and the alternative system in a 15-minute tutorial. Participants were then randomly assigned to Group A or B each sized six. Datasets Iris and Driving are explored sequentially, and to compare systems and mitigate learning effects, Group A used Voyager first and then Dowsing, while Group B used Dowsing first and then Voyager. During exploration, we recorded time taken, charts created (Fig. 5), and screens. Finally, we interviewed participants for subjective feedback. The full study lasted approximately 1 h.

Results. Quantitative analysis showed that participants spent less time and created fewer visualizations with Dowsing versus the alternative (Fig. 6). On average, the participants took 14.7 min per dataset (SD=3.3) with the alternative system while took 11.9 (SD = 2.7) with Dowsing. The Driving dataset required less time for exploration on both systems since given fewer fields. Notably, there was a significant difference in the total time spent exploring datasets between Dowsing and the alternative system, as well as in the time spent specifically exploring the Iris dataset, with p-values of 0.046 and 0.034, respectively. These statistical results suggest that Dowsing allowed for more efficient data exploration compared to the alternative system, particularly in the case of the Iris dataset. On Iris, participants made 7.8 charts with Dowsing (SD = 2.1) versus 9.3 with the alternative (SD = 3.6), and it's similar on Driving for 7.8 charts with Dowsing vs. 7.3 charts with the alternative. In interviews, most participants reported Dowsing provided more helpful recommendations than the alternative system. P5 noted Dowsing's recommendations went beyond the current view, enriching exploration diversity. P2 said *the recommended charts were highly relevant, enabling deep target analysis*. P10 found *Dowsing initially gave recommendations like Voyager, but with continued use, more useful charts emerged, improving efficiency*. We noticed some participants said the alternative system recommended multiple similar visualizations, which may have facilitated more charts but not always effective for exploration. P11 stated *Voyager drove deeper analysis while Dowsing broadened it*. P1 said, "*I noticed Voyager always gives more results than Dowsing, so I have spent more time checking them, but many were very similar and gave insights with no surprise.*" Many participants praised Dowsing's controllability. P6 said, "*Unlike Voyager, I can adjust the active task through the panel in Dowsing, which is*

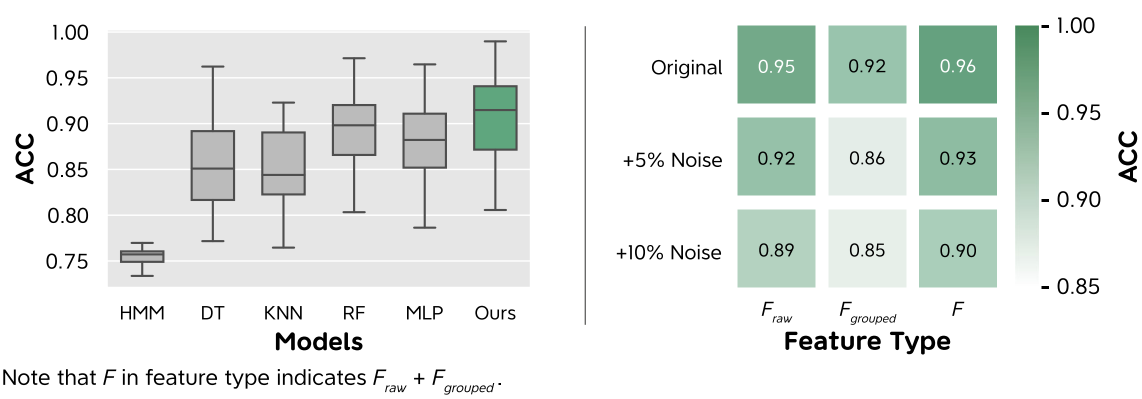

Fig. 4 Quantitative evaluation of the task prediction model. The left side compares our Bi-LSTM-based model against baselines, demonstrating its superior performance. The right side shows the feature ablation study, indicating that both the metric performance and robustness are improved by concatenating raw and grouped features

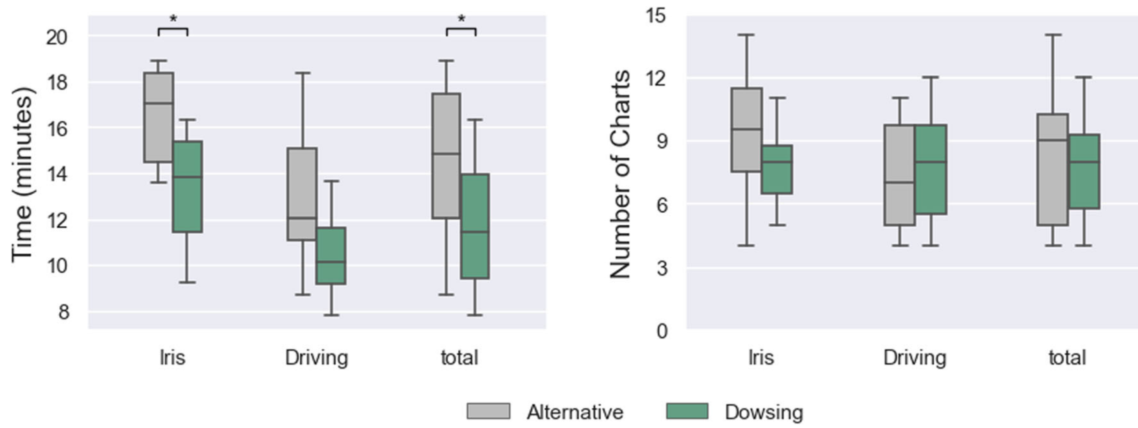

Fig. 5 Quantitative analysis results of user study 1. The left displays the discrepancy in time spent by participants to fulfill data exploration by utilizing Dowsing and the alternative system, and the right exhibits the variance in the quantity of charts generated by participants during data exploration. Data groups exhibiting significant differences ($p < 0.05$) are marked with connecting lines and asterisks (*)

a very refreshing feeling!" P12 felt *Dowsing's grid-based canvas helped in organizing charts during exploration to synthesize insights*. In conclusion, quantitative and qualitative evaluations in this comparative study demonstrate Dowsing improves data exploration efficiency over the alternative system.

6.2.2 User study 2: behavioral patterns and subjective evaluation

We conducted a semi-structured user study to investigate two key questions: (1) Are all the features offered by Dowsing fully utilized by users? and (2) Are users satisfied with the user experience with Dowsing? These questions validate the rationality and practicality of Dowsing as a mixed-initiative system for data exploration.

Participants. We recruited fifteen participants (twelve males and three females) with experience in data analysis who were either developers working in internet companies or senior-grade undergraduates. Five of the participants had no prior experience in visualization development, while the others had up to six years of experience. The participants in this study were distinct from those in User Study 1 and are labeled as U1-U15 for differentiation.

Datasets. We chose three real-world datasets, namely Weather, Movies, and Birdstrikes, each containing over 10 data fields and 2000 records.

Procedure. The study began with a 10-minute recorded tutorial of Dowsing using the Weather dataset. Participants were introduced with the study procedure and were allowed to seek help. They then completed an initial USE questionnaire on usability and learnability. Dataset order was balanced via a random number table. In two 20-minute exploration rounds with Dowsing, participants constructed visualizations to address given needs, recording satisfied visualizations and their purpose. After each round, participants completed another USE questionnaire and rested up to 5 min. Finally, we interviewed participants about their usage experience, problems encountered, interface comments, and improvement suggestions.

Results. Participants responded positively across the four subjective measures (Fig. 6; raw data in Appendix 7). Interviews revealed that participants particularly favored Dowsing's clean, controllable interface. Analyzing their comments and study records, we identified three common usage patterns:

- *Task-driven insight finding*: In this pattern, users initially edit tasks and subsequently explore the dataset through recommendations. Seven of the fifteen participants had similar comments and behaviors. U4 typified this in saying, "*I first conceived the problem I wanted to solve.*" U7 remarked, "*After viewing the raw data, I assumed interesting fields and tasks. Rather than an overview, I learned about the data by completing expected tasks.*" This validates our motivation for task-aware visualization recommendations. Notably, this behavior always occurred early on exploration, except for one experienced participant who consistently followed this pattern.
- *Speedup the process of exploration*. The characteristic of this pattern is frequently checking recommendations and predicted tasks. Among the fifteen participants, twelve had such behaviors and comments. When providing recommendations only, they felt confuse and took time to understand the

visualizations and relationships between fields (G1). U8 remarked, "*First I choose x- and y-axes, seeing if the chart satisfies my needs. If not, I check recommendations and change settings.*" Participants felt Dowsing's explicit task and history changes helped them recall exploration purposes (G2).

- *Capturing the overlooked data*. This pattern is typically characterized by participants examining recommendations and tasks after they have almost completed their exploration. Ten of the fifteen participants switched all task tags and viewed matching recommendations after constructing visualizations, aligning with G4. As U9 said, "*After completing the visualization, I get inspiration here (the agency controller)*". All participants utilized the canvas to organize their multi-view visualizations, meeting goal G5. U15 said, "B*y organizing views with similar fields together, I can compare insights across (views) and observe new field combinations.*"

Overall, Dowsing received positive feedback, with the usage aligning with expectations. Review of screen recordings showed that all users utilized the control panel, suggestion panel, and canvas. However, only four participants used the history panel to review visualization edits, likely because of the limited number of edits when creating simple visualizations. U6 said, "*I typically add channels one by one until an insight is formed; I rarely compare history,*" which prompted us to modify the layout of the history panel to allow for automatic folding. Additionally, U2 suggested manual control over field preferences, which we implemented by modifying the control panel later.

7 Discussion

In this section, we discuss the key findings and implications of our study, focusing on the controllability of Dowsing through its modular architecture, the potential for generalizable analytical task prediction, and the support for multi-view visualizations.

Human in loop. Through expert interviews, we identified controllability as a key design goal. Dowsing realizes controllability via a multi-stage workflow, decomposing the system into the stages of a human–computer mixed agency and a visualization recommender. This modular architecture empowers users to independently steer each process's output. For instance, users can manually adjust the analytical task when its predictions mismatch expectations. By manually managing field preferences, users also focus analyses on pertinent data subsets. These features are confirmed in user studies with extensive utilizations. Enabling users to tweak recommendation algorithms to meet overlooked needs merits consideration in future visualization systems.

Pursuing generalizable analytical task prediction. We currently employ an LSTM-based model to dynamically predict users' analytical tasks from their action logs across thirty operation types and six tasks. User studies have evidenced its promising task prediction capabilities during tabular data exploration, which motivates us to expand task prediction to novel scenarios like weather and genomics data analysis. However, scenario-specific actions and tasks may necessitate custom models and substantial dataset collection. Rather than such time-intensive projects, we advocate identifying universal user action representations to construct large-scale, cross-scenario log datasets. Such representations can enable building widely generalizable, transferable models to unlock analytical task prediction's full potential.

Supporting multi-view visualizations. Dowsing aims to enable fast generation of visualizations for data analysis. Expert interviews highlighted the value of arranging charts into dashboards. This inspired

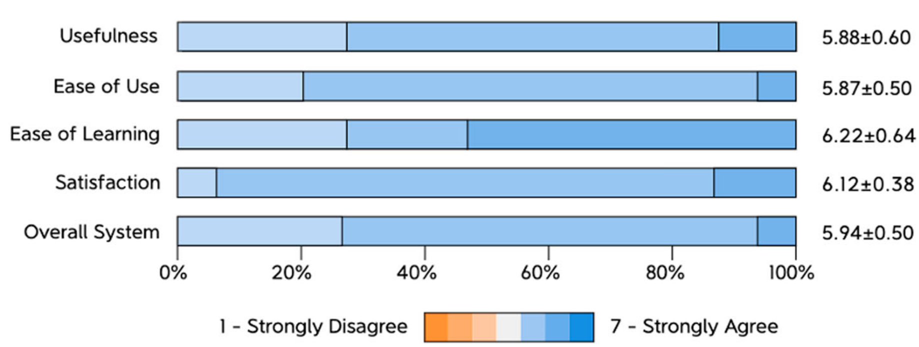

Fig. 6 Subjective evaluation of participants on Dowsing

Dowsing's grid canvas for layout. User studies confirm the grid canvas's utility. Moreover, though not exclusively designed, Dowsing also facilitates the creation of multi-view visualizations. As U14 reported, "Recommendations relate to the locked task, so themes stay consistent when I make multiple visualizations in a sequence." This underscores the feasibility of intermediate task representations to coherently guide multi-view recommendations. Challenges remain, like ensuring visual encoding consistency, but Dowsing represents progress.

Overcoming Data Scale Limitations. The user action data were gathered within a controlled workshop environment, thereby covering only a limited array of action types, user numbers, and chart types. These limitations narrow the design space supported by Dowsing. Although Dowsing's recommendation workflow can be easily expanded to adapt to a greater variety of charts and action types, collecting action data with task labels in visual data exploration remains a high-cost task. This is because it necessitates action data from users with expertise in visualization, along with corresponding natural language annotations, which must then be manually translated into task labels. We look forward to the potential application of generative AI technologies, such as large language models (LLMs), in the data collection process. For instance, employing language models to devise appropriate operational steps for specific datasets and analytical tasks could lead to the automatic generation of matched action sequence-task label data. Additionally, these technologies could facilitate the automatic assignment of suitable task labels to natural language annotations provided by users.

8 Conclusion

We present Dowsing, a dynamic visualization recommendation system for data exploration. Dowsing analyzes action logs to recommend visualizations aligning with users' requirements derived from predicted analytical tasks and detected field preferences. Its recommendation workflow integrates LSTM-based task prediction, Chi-square field preference detection, and answer set programming-based visualization recommendation. Quantitative experiments prove that the task prediction model's superiority on accuracy and robustness against noises. Finally, user studies evaluate Dowsing's overall effectiveness.

However, user studies and critical discussion reveal limitations we aim to address. To improve cold start performance, we will incorporate data insight generation to spawn predefined analytical tasks. We also plan to collect diverse domain action data to boost generalizability of task prediction. Finally, integrating visualization anchoring techniques Lin et al. (2020) into the recommender could enhance visual channel consistency for multi-view analysis.

Funding This research is partially supported by National Natural Science Foundation of China (62172289) and the School-City Strategic Cooperation Project (2021CDSN-13).

References

Albers D, Correll M, Gleicher M (2014) Task-driven evaluation of aggregation in time series visualization. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 551–560. ACM, Toronto Ontario Canada . https://doi.org/10.1145/2556288.2557200

Amar R, Eagan J, Stasko J (2005) Low-level components of analytic activity in information visualization. In: IEEE symposium on information visualization, 2005. INFOVIS 2005., pp. 111–117

Battle L, Heer J (2019) Characterizing Exploratory Visual Analysis: A Literature Review and Evaluation of Analytic Provenance in Tableau. Computer Graphics Forum 38(3):145–159. https://doi.org/10.1111/cgf.13678

Brehmer M, Munzner T (2013) A multi-level typology of abstract visualization tasks. IEEE Trans Visual Comput Graphics 19(12):2376–2385

Breiman L (2001) Random forests. Mach Learn 45:5–32

Brown ET, Ottley A, Zhao H, Lin Q, Souvenir R, Endert A, Chang R (2014) Finding Waldo: learning about users from their interactions. IEEE Trans Visual Comput Graphics 20(12):1663–1672. https://doi.org/10.1109/TVCG.2014.2346575

Cao Y-R, Xiao-Han Li, Pan J-Y, Lin W-C (2022) VisGuide: user-Oriented Recommendations for Data Event Extraction. In: CHI conference on human factors in computing systems, pp. 1–13. ACM, New Orleans LA USA. https://doi.org/10.1145/3491102.3517648

Chen X, Zeng W, Lin Y, AI-maneea HM, Roberts J, Chang R (2021) Composition and configuration patterns in multiple-view visualizations. IEEE Trans Visual Comput Graphics 27(2):1514–1524

Collins C, Andrienko N, Schreck T, Yang J, Choo J, Engelke U, Jena A, Dwyer T (2018) Guidance in the human-machine analytics process. Visual Inform 2(3):166–180

Dabek F, Caban JJ (2017) A grammar-based approach for modeling user interactions and generating suggestions during the data exploration process. IEEE Trans Visual Comput Graphics 23(1):41–50. https://doi.org/10.1109/TVCG.2016.2598471

de Oliveira MCF, Levkowitz H (2003) From visual data exploration to visual data mining: A survey. IEEE Trans Visual Comput Graphics 9(3):378–394. https://doi.org/10.1109/TVCG.2003.1207445

Gotz D, Wen Z (2009) Behavior-driven visualization recommendation. In: Proceedings of the 14th international conference on intelligent user interfaces, pp. 315–324. ACM, Sanibel Island Florida USA

Gotz D, Zhou MX (2008) An empirical study of user interaction behavior during visual analysis. IBM Research RC24525 (W0803-127)

Guo H, Gomez SR, Ziemkiewicz C, Laidlaw DH (2016) A case study using visualization interaction logs and insight metrics to understand how analysts arrive at insights. IEEE Trans Visual Comput Graphics 22(1):51–60. https://doi.org/10.1109/TVCG.2015.2467613

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kim Y, Heer J (2018) Assessing effects of task and data distribution on the effectiveness of visual encodings. Computer Graphics Forum 37(3):157–167. https://doi.org/10.1111/cgf.13409

Li H, Wang Y, Zhang S, Song Y, Qu H (2022) KG4Vis: a knowledge graph-based approach for visualization recommendation. IEEE Trans Visual Comput Graphics 28(1):195–205. https://doi.org/10.1109/TVCG.2021.3114863

Li Y, Qi Y, Shi Y, Chen Q, Cao N, Chen S (2022) Diverse interaction recommendation for public users exploring multi-view visualization using deep learning. IEEE Trans Visual Comput Graphics. https://doi.org/10.1109/TVCG.2022.3209461

Lin H, Moritz D, Heer J (2020) Dziban: Balancing Agency & Automation in Visualization Design via Anchored Recommendations. In: Proceedings of the 2020 CHI Conference on human factors in computing systems, pp. 1–12. ACM, Honolulu HI USA

Luo X, Yuan Y, Zhang K, Xia J, Zhou Z, Chang L, Gu T (2019) Enhancing statistical charts: toward better data visualization and analysis. J Visualization 22:819–832

Luo Y, Qin X, Tang N, Li G (2018) DeepEye: towards automatic data visualization. In: 2018 IEEE 34th international conference on data engineering (ICDE). IEEE, Paris, pp. 101–112

Moritz D, Wang C, Nelson GL, Lin H, Smith AM, Howe B, Heer J (2019) Formalizing visualization design knowledge as constraints: actionable and extensible models in Draco. IEEE Trans Visual Comput Graphics 25(1):438–448

Narechania A, Srinivasan A, Stasko J (2021) NL4DV: a toolkit for generating analytic specifications for data visualization from natural language queries. IEEE Trans Visual Comput Graphics 27(2):369–379

Nguyen QV, Miller N, Arness D, Huang W, Huang ML, Simoff S (2020) Evaluation on interactive visualization data with scatterplots. Visual Inform 4(4):1–10

Ojo F, Rossi RA, Hoffswell J, Guo S, Du F, Kim S, Xiao C, Koh E (2022) VisGNN: Personalized Visualization Recommendationvia Graph Neural Networks. In: Proceedings of the ACM web conference 2022, pp. 2810–2818. ACM, Virtual Event, Lyon France. https://doi.org/10.1145/3485447.3512001

Peterson LE (2009) K-nearest neighbor. Scholarpedia 4(2):1883

Popescu M-C, Balas VE, Perescu-Popescu L, Mastorakis N (2009) Multilayer perceptron and neural networks. WSEAS Trans Circ Syst 8(7):579–588

Qian X, Rossi RA, Du F, Kim S, Koh E, Malik S, Lee TY, Ahmed NK (2022) Personalized visualization recommendation. ACM Trans Web (TWEB) 16(3):1–47

Qin X, Luo Y, Tang N, Li G (2020) Making data visualization more efficient and effective: a survey. VLDB J 29(1):93–117. https://doi.org/10.1007/s00778-019-00588-3

Quadri GJ, Rosen P (2022) A survey of perception-based visualization studies by task. IEEE Trans Visual Comput Graphics 28(12):5026–5048. https://doi.org/10.1109/TVCG.2021.3098240

Rabiner L, Juang B (1986) An introduction to hidden Markov models. IEEE ASSP Mag 3(1):4–16. https://doi.org/10.1109/MASSP.1986.1165342

Rind A, Aigner W, Wagner M, Miksch S, Lammarsch T (2016) Task Cube: a three-dimensional conceptual space of user tasks in visualization design and evaluation. Inf Vis 15(4):288–300

Roberts JC (2007) State of the Art: coordinated & multiple views in exploratory visualization. In: Fifth international conference on coordinated and multiple views in exploratory visualization (CMV 2007), pp. 61–71. https://doi.org/10.1109/CMV.2007.20

Roth SF, Mattis J (1990) Data characterization for intelligent graphics presentation. In: Proceedings of the SIGCHI conference on human factors in computing systems empowering people - CHI ’90, pp. 193–200. ACM Press, Seattle, Washington, USA. https://doi.org/10.1145/97243.97273

Rubab S, Tang J, Wu Y (2021) Examining interaction techniques in data visualization authoring tools from the perspective of goals and human cognition: a survey. J Visualiz 24:397–418

Saket B, Endert A, Demiralp Ç (2019) Task-Based Effectiveness of Basic Visualizations. IEEE Trans Visual Comput Graphics 25(7):2505–2512. https://doi.org/10.1109/TVCG.2018.2829750

Sarikaya A, Gleicher M (2018) Scatterplots: tasks, data, and designs. IEEE Trans Visual Comput Graphics 24(1):402–412

Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE Trans Signal Process 45(11):2673–2681. https://doi.org/10.1109/78.650093

Shen L, Shen E, Tai Z, Song Y, Wang J (2021) TaskVis: task-oriented Visualization Recommendation. EuroVis 2021 - Short Papers, 5

Shirato G, Andrienko N, Andrienko G (2023) Exploring and visualizing temporal relations in multivariate time series. Visual Inform 7(4):57–72. https://doi.org/10.1016/j.visinf.2023.09.001

Shi D, Shi Y, Xu X, Chen N, Fu S, Wu H, Cao N (2019) Task-oriented optimal sequencing of visualization charts. In: 2019 IEEE visualization in data science (VDS), pp. 58–66

Song Y-Y, Ying L (2015) Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 27(2):130

Soni P, de Runz C, Bouali F, Venturini G (2024) A survey on automatic dashboard recommendation systems. Visual Informatics. https://doi.org/10.1016/j.visinf.2024.01.002

Stoiber C, Ceneda D, Wagner M, Schetinger V, Gschwandtner T, Streit M, Miksch S, Aigner W (2022) Perspectives of visualization onboarding and guidance in VA. Visual Inform 6(1):68–83. https://doi.org/10.1016/j.visinf.2022.02.005

Vladimir L (2008) What is answer set programming. In: Proc. 23rd AAAI conference on artificial intelligence, vol. 8, pp. 1594–1597

Wongsuphasawat K, Moritz D, Anand A, Mackinlay J, Howe B, Heer J (2016) Towards a general-purpose query language for visualization recommendation. In: Proceedings of the workshop on human-in-the-loop data analytics, pp. 1–6. ACM, San Francisco California. https://doi.org/10.1145/2939502.2939506

Wongsuphasawat K, Qu Z, Moritz D, Chang R, Ouk F, Anand A, Mackinlay J, Howe B, Heer J (2017) Voyager 2: augmenting visual analysis with partial view specifications. In: Proceedings of the 2017 CHI conference on human factors in computing systems, pp. 2648–2659. ACM, Denver Colorado USA

Wu A, Wang Y, Zhou M, He X, Zhang H, Qu H, Zhang D (2021) Multivision: designing analytical dashboards with deep learning based recommendation. IEEE Trans Visual Comput Graphics 28(1):162–172. https://doi.org/10.1109/tvcg.2021.3114826

Wu A, Wang Y, Shu X, Moritz D, Cui W, Zhang H, Zhang D, Qu H (2022) AI4VIS: survey on artificial intelligence approaches for data visualization. IEEE Trans Visual Comput Graphics 28(12):5049–5070. https://doi.org/10.1109/TVCG.2021.3099002

Xu K, Attfield S, Jankun-Kelly TJ, Wheat A, Nguyen PH, Selvaraj N (2015) Analytic provenance for sensemaking: a research agenda. IEEE Comput Graphics Appl 35(3):56–64. https://doi.org/10.1109/MCG.2015.50

Yu D, Ian O, Jie L, Xiaoru Y, Vinh NQ (2023) User-centered visual explorer of in-process comparison in spatiotemporal space. J Visualization 26(2):403–421. https://doi.org/10.1007/s12650-022-00882-3

Zhang X, Ge X, Chrysanthis PK, Sharaf MA (2021) ViewSeeker: an interactive view recommendation framework. Big Data Research 25:100238. https://doi.org/10.1016/j.bdr.2021.100238

Zhou Y, Meng X, Wu Y, Tang T, Wang Y, Wu Y (2023) An intelligent approach to automatically discovering visual insights. J Visualiz 26(3):705–722. https://doi.org/10.1007/s12650-022-00894-z

Zhu S, Sun G, Jiang Q, Zha M, Liang R (2020) A survey on automatic infographics and visualization recommendations. Visual Inform 4(3):24–40. https://doi.org/10.1016/j.visinf.2020.07.002

Zhu Z, Shen Y, Zhu S, Zhang G, Liang R, Sun G (2023) Towards better pattern enhancement in temporal evolving set visualization. J Visualiz 26(3):611–629. https://doi.org/10.1007/s12650-022-00896-x