Enhancing Table-to-Text Generation: A Focus on Subjectivity and Objectivity with the TaTS Dataset

Anonymous ACL submission

Abstract

Table-to-text generation, a long-standing challenge in natural language generation, has remained unexplored through the lens of subjectivity. Subjectivity here encompasses the comprehension of information derived from the table that cannot be described solely by objective data. To ascertain the relevance and social applications of this work, we conduct a public survey involving relevant people. The survey results unequivocally conclude the significance of this research. Given the absence of pre-existing datasets, we introduce the **TaTS dataset**. We perform the task using various sequence-to-sequence models on the flattened table and analyze the results from a qualitative perspective to ensure the capture of subjectivity. To the best of our knowledge, this is the first kind of dataset on tables with subjectivity included.

1 Introduction

In the contemporary era of big data, humongous volumes of information are being generated and archived in numerous different formats. Among these, tables stand out as a very useful technique for structured data storage. Using the potential of natural language to comprehend tabular data holds the promise of enhancing human efficiency across diverse applications. Text generation from tables has seen some contributions previously (See Section 2). These efforts have predominantly centered on either the generation of text from relatively simple tabular structures or the translation of numerical data into natural language, devoid of nuanced interpretation. In this research, we propose a distinctive point of view to look at the table-to-text generation problem statement.

1.1 Problem Statement

We introduce a novel problem statement with the aim of text generation from tabular data, devoted to infusion of subjectivity into the generated text. Subjectivity, in the context, refers to the nuanced interpretation of the data, beyond the realm of numeric or objective representation.

Table:

Particulars	*2019*	*2020*	*2021*	*2022*	*2023*
Revenue	2,462	2,365	3,634	4,316	4,197
Expenses	2,084	2,080	2,088	2,701	2,955
EBITDA	918.7	819.37	1,936	2,139	1,896
EBIT	894.75	779.66	1,889	2,091	1,838
Profit Before Tax	377.9	285.19	1,458	1,616	1,242
Net Profit	285.25	215.4	1,203	1,311	932.82
Financial Ratios		None	None	None	None
Operating Profit Margin (%)	37.32	34.64	53.28	49.55	45.18
Net Profit Margin (%)	11.59	9.11	33.10	30.37	22.23
Earnings Per Share (Diluted)	19.67	12.2	84.21	88.38	62.66
Dividends Per Share	8.5	4	10	10	10

Reference Text:

The company's income statement shows its financial performance in the past 5 years with the revenue showing a **minor dip** in 2020 touching 2,365 Cr. Rs. and then **increasing steadily** until 2022 to 4,316 Cr. Rs. The year 2023 has witnessed another **slight decline** in revenues, throwing light on the company's slow growth. Expenses have increased by **greater margins** 2022 onwards reaching 2,701 Cr. Rs. and thus, resulting in decreased profit margin. Positive EPS in 2021-2022 **signals profitability** per share which is **crucial for shareholders**. The company's commitment to maintain a **constant** value of 10.00 Rs. of dividends per share demonstrates an **overall balance** in the company. Thus, the company statistics reflect a **steady growth** with recent slowdown which hints at a **decline of performance** and operational efficiency. Investors with future prospects with this company should hold cautious **optimism**.

Figure 1: Generating text with subjectivity from a table: The table contains the income statement report of a company over 5 years. The reference text below describes the tabular information where the bold phrases refer to the infused subjectivity. This subjectivity is the interpretation of the data from a human perspective. (Due to space constraints, shown in figure 3)

The table depicted in Figure 1 presents a company's income statement table with 10 features such as Revenue, Expenses, EBIDTA, EBIT, Profit before taxes, Net profit, Financial ratios, Operating profit margin, Net profit margin, Earnings per

share and Dividends per share respectively. Within the accompanying narrative of the table, several expressions are employed, which do not find direct representation in the objective data. Instead, these phrases refer to the interpretation of the underlying sentiment associated with the data.

The phrase ***minor dip*** comes from the idea that having a downfall of 200 Cr. rs. is **small** compared to a growth of 2000 Cr. rs. over the last few years. Whereas, ***signals profitability*** comes from the understanding that positive EPS refers to positive growth.

1.2 Problem Formulation

The problem of table-to-text generation takes a table T as input where $T = \{t_{1,1}, .., t_{i,j}, .., t_{m,n}\}$ where $t_{i,j}$ refers to the token of ith row and jth column with at most m rows and n columns. A pre-processed input document Y is produced from the table where Y can be written as, $Y = y_1 y_2 .. y_l$. Here y_i is the ith token in Y which has l tokens and $l \geq mn$. In an auto-regressive approach text $S = s_1 s_2 ... s_k$ is generated from Y which can be defined as, $s_i = argmax P(s_i | Y, s_1, s_2, ..., s_{i-1}; \theta)$ where θ is model parameters.

1.3 Motivation

Text generation from tables can be useful in various social applications where a mere description of the numerical values proves to be insufficient.

- News or Blog writings on a Tournament points table, Pricing table, or Voting results table.
- Reports on a Match summary, Business details, Sales detail of a company.
- Explaining Healthcare reports, Weather reports, or understanding Legal documents, etc.

To establish the practical utility of this research, a comprehensive survey is undertaken. The survey encompassed 4 distinct tables, each accompanied by corresponding reference texts with subjective phrases. These tables spanned diverse genres, such as points tables, pricing tables, real-estate property data, and weather forecast tables.

Approximately **2,400** participants actively engaged in the survey. In light of the statistics from figure 5 (see A.2), the survey unequivocally underscores the profound impact and efficacy of this research in enhancing the everyday experiences of the common people.

1.4 Our Contributions

Our contributions are:

1. The formulation of a problem statement of subjective text generation from tables, enriched with social significance.
2. The creation of the **TaTS** dataset for this task comprising **3077 table instances** sourced from sports, financial statements, and weather forecast domains.
3. The detailed analysis of the generated text revealing **87%** subjectivity being captured. A new table linearization technique has been used to flatten the tables for sequence-to-sequence models.

2 Related Works

The table-to-text problem contains several different approaches based on the type of table. The **Wikibio-infobox** problem (Lebret et al., 2016) involves generating textual descriptions from tables with just one column extracted from Wikipedia infoboxes. Versatile neural language models and encoders were introduced to tackle this challenge by Liu et al. (2017) and Rebuffel et al. (2022).

The **ToTTo** dataset (Parikh et al., 2020a) represents another prominent objective of text generation from a single highlighted row of the table. Gehrmann et al. (2021) harnessed the T5 model to address this challenge. Subsequently, sequence-to-sequence techniques and structure-aware frameworks were introduced by Parikh et al. (2020b) and Wang et al. (2022). The current state-of-the-art on this dataset has been held by Kale and Rastogi (2020), who leveraged the T5 model with 3 billion parameters to achieve the highest scores. A very similarly aligned problem, **Wiki-table-to-text**, has also been worked on by Chen et al. (2021).

Among the previously discussed problem statements, none encompass the task of comprehending complicated tables with multiple rows and columns. The **Rotowire** dataset (Wiseman et al., 2017) is expressly designed to address this challenge. The concepts of macro-planning, content selection, and planning have been significant contributions to this field by Puduppully and Lapata (2021), Puduppully et al. (2019). Li et al. (2021) achieved state-of-the-art results by introducing a Record encoder, a Reasoning module, and a Decoder equipped with Dual attention. Additionally, alternative approaches have also been explored by Rebuffel et al. (2019), (Gong et al., 2019) and Choi et al. (2021). It is worth noting that the RotoWire tables exhibit a notable degree of similarity to the tables of our interest. The

flattened tables derived from data instances within RotoWire can contain up to 17,000 tokens. Most of the current language model architectures do not support such a huge context length. Thus, using pre-trained models for this task is very challenging. While training a model from the ground up may address this challenge, there are concerns regarding its comprehensive understanding of the language. Furthermore, the utility of such a model in processing tables from other sports domains or tables of completely different genres is further questionable.

3 Dataset

We contribute to building a novel dataset for this problem statement which is named as **TaTS Dataset (Table to Text with Subjectivity)**.

3.1 Data Collection and Filtration

The dataset encompasses three types of tables, such as financial statements, weather forecast reports, and sports tables. We collected the tables using various web-based platforms that are entitled to showcase different tables. Five sites were selected as sources of the tables such as Groww, Indian Meteorological Dept., ESPN Cric Info, IPL, and Goal (see A.3 for sources). The tables are scraped using Python Programming language. A total of 3077 tables were collected from the sources by web scraping. The complete dataset statistics can be found in the table 3 in the Appendix section (see A.4). The collected data is filtered manually. Tables with less number rows and columns that threshold are removed. Also, data instances having very low cell values are discarded.

3.2 Annotation Details

We employ four annotators for this task who are proficient in the English language. Two of the annotators are graduates of Computer science and engineering while the others are PhD students and postgraduates in English literature respectively.

They were provided with detailed annotation guidelines with rules, objectives, and multiple examples. Each instance in the set was annotated only once. The distribution among the annotators was 700, 700, 850, and 827 instances. An example annotation is shown in figure **??** (See A.5). 30 samples were annotated by all 4 annotators which are used to compute the Bleu-4 score for **Inter-Annotator Agreement** (IAA). The pair-wise average B-4 score is shown in section A.5. The annotation may differ in the use of different words having the same intent or order of description from the table. However, the Bleu score signifies the similarity in the instances.

4 Experimental Approaches

We performed experiments on a subset of the mentioned TaTS dataset using sequence-to-sequence models. The sports genre tables are experimented on for now and due to the relatively small size of the set, we choose pre-trained models for our experiments.

4.1 Linearization of Tables

Transformer-based sequence-to-sequence models are capable of understanding linearized information with long context. A closer examination of the reference texts reveals that the textual content contains two distinct types of information, namely, row-wise and column-wise comparison. For example, the income-statement table depicts phrases that express the relative growth of the company in time which requires row-wise comparison and relation between revenue-EBIDTA or other features which requires column-wise comparison. Hence, the application of both row-major flattening and column-major flattening techniques is required. Thus, row-major and column-major flattenings are generated and concatenated with delimiter tokens (examples shown in A.6).

4.2 Sequence-to-sequence Model

We use the T5 (Raffel et al., 2020) and BART (Lewis et al., 2020) models to perform the task. T5 is a transformer-based sequence-to-sequence learner. Multiple versions of the T5 model such as T5-small, base, and large are used for the experiments. The overall architecture with flattening and the model usage is shown in figure 2.

4.3 Experimental Setup

The T5 models are fine-tuned with the prefix *Describe the table briefly:* with a batch size of 32 for 100 epochs in each case. The batch sizes are varied to check for the best case. For T5 models we use a learning rate (lr) of $1e-3$ and a batch size of 16 (see details in A.7). A train-valid-test split of 80%-10%-10% is used.

5 Results

We measured the generation of the fine-tuned models through automatic and human evaluation methods.

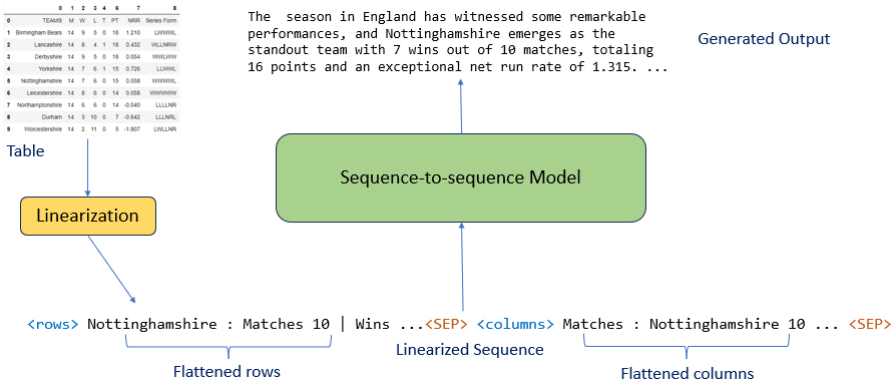

Figure 2: Architecture: Tables are flattened through linearization which is then appended with delimiter tokens. This sequence acts as input to various sequence-to-sequence models which generates the expected output.

5.1 Quantitative Analysis

Model	Bleu-4	Meteor	Rouge-L
T5-small	10.58	13.77	8.13
T5-base	15.11	19.14	11.90
T5-large	21.09	24.54	11.24
BART	10.42	12.33	9.54

Table 1: Average Bleu-4 (Papineni et al., 2002), Meteor (Banerjee and Lavie, 2005), and Rouge-L (Lin, 2004) score on seen and unseen text by various models on the sports tables of **TaTS Dataset**.

The best Bleu-4 metric score achieved is around **21.09**. The RotoWire problem (see section 2) has a state-of-the-art bleu score of 19.43 and is a very similarly aligned problem to our task without subjectivity. Hence, we interpret our bleu score to be decent. The RotoWire dataset, consisting of large sports tables with multiple columns and rows which are to be described in the generated text, has a state-of-the-art score of around 19.43 Bleu score. Our task is similar to the RotoWire problem, hence, we assume the Bleu score to be decent.

With long reference texts, an n-gram overlap cannot justify the model's efficiency. Moreover, the subjectivity infusion can also not be verified from these metrics. Hence, we perform a human evaluation to understand the efficiency of the model.

5.2 Qualitative Analysis

Figure 7 shows an example of generated output by the model (see in section A.8). The generated instance signifies sound capture of subjective phrases at the right places. However, it lacks coverage, and some wrong information is also generated. Further fine-tuning or the use of a heavier model may reduce this problem.

We perform a human evaluation to asses the generated results. 30 randomized samples of generated instances were provided to the annotators to assess over *coherence (Coh.)*, *coverage (Cov.) and Accuracy(Acc.)*, and *subjectivity capture (Sub. Cap.)*, shown in table 2. The details about guidelines of human evaluation are in section A.8.

	Coh.	Cov. & Acc.	Sub. Cap.
A	8.4	6.3	9.1
B	7.9	5.2	8.7
C	8	5.7	8.4

Table 2: Human evaluation: the average score of all 20 generated samples over the given features.

From the table 2 it is understandable the generations can capture the subjectivity fruitfully and understand the semantics of the data. However, the coverage and accuracy having a low score signifies the room for improvement in this.

6 Conclusion and Future Work

In this paper, we present a novel perspective on the domain of table-to-text generation by introducing the element of subjectivity, an unexplored dimension of this field for several different table genres. To facilitate our goal, we curate the **TaTS** Dataset, develop a linearization technique, and conduct fine-tuning experiments to assess the effectiveness of various models in capturing subjectivity. The qualitative evaluations indicate that the models exhibit a robust understanding of word knowledge, with subjective phrases being appropriately reflected in the generated text. However, there is ample room for further exploration. Expanding the task to encompass a larger and more diverse dataset, comprising tables from other genres, would be a valuable approach for future research. Moreover, the increase in metrics, with the current results serving as a benchmark, holds the potential to advance the state of the art in this field.

7 Limitations

Our dataset contains tables from three genres. Subjectivity in the same type of table can be influenced by similar semantics, which may not be the case for a more diverse dataset. Hence, building a robust system to handle all kinds of tables will be more feasible with an even more diverse or large dataset. Moreover, we developed a table flattening technique to fit tables in sequence-to-sequence models which are trained on normal text. For prospects, we propose ideas to fuse pre-trained table encoders with decoders and fine-tune our dataset to get even better metric scores. Also, a necessity of measuring the generations based on factual consistency lies ahead in sight.

8 Ethics Statement

All of our collected data were present in open-source mediums and did not contain personal, restricted, or illegal data. None of the generated texts or annotated texts were intended to promote or derogate any team or entity.

References

Satanjeev Banerjee and Alon Lavie. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 65–72, Ann Arbor, Michigan. Association for Computational Linguistics.

Mingda Chen, Sam Wiseman, and Kevin Gimpel. 2021. WikiTableT: A large-scale data-to-text dataset for generating Wikipedia article sections. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 193–209, Online. Association for Computational Linguistics.

Sanghyuk Choi, Jeong in Hwang, Hyungjong Noh, and Yeonsoo Lee. 2021. May the force be with your copy mechanism: Enhanced supervised-copy method for natural language generation.

Sebastian Gehrmann, Tosin Adewumi, Karmanya Aggarwal, Pawan Sasanka Ammanamanchi, Anuoluwapo Aremu, Antoine Bosselut, Khyathi Raghavi Chandu, Miruna-Adriana Clinciu, Dipanjan Das, Kaustubh Dhole, Wanyu Du, Esin Durmus, Ondřej Dušek, Chris Chinenye Emezue, Varun Gangal, Cristina Garbacea, Tatsunori Hashimoto, Yufang Hou, Yacine Jernite, Harsh Jhamtani, Yangfeng Ji, Shailza Jolly, Mihir Kale, Dhruv Kumar, Faisal Ladhak, Aman Madaan, Mounica Maddela, Khyati Mahajan, Saad Mahamood, Bodhisattwa Prasad Majumder, Pedro Henrique Martins, Angelina McMillan-Major, Simon Mille, Emiel van Miltenburg, Moin Nadeem, Shashi Narayan, Vitaly Nikolaev, Andre Niyongabo Rubungo, Salomey Osei, Ankur Parikh, Laura Perez-Beltrachini, Niranjan Ramesh Rao, Vikas Raunak, Juan Diego Rodriguez, Sashank Santhanam, João Sedoc, Thibault Sellam, Samira Shaikh, Anastasia Shimorina, Marco Antonio Sobrevilla Cabezudo, Hendrik Strobelt, Nishant Subramani, Wei Xu, Diyi Yang, Akhila Yerukola, and Jiawei Zhou. 2021. The GEM benchmark: Natural language generation, its evaluation and metrics. In *Proceedings of the 1st Workshop on Natural Language Generation, Evaluation, and Metrics (GEM 2021)*, pages 96–120, Online. Association for Computational Linguistics.

Li Gong, Josep Crego, and Jean Senellart. 2019. Enhanced transformer model for data-to-text generation. In *Proceedings of the 3rd Workshop on Neural Generation and Translation*, pages 148–156, Hong Kong. Association for Computational Linguistics.

Mihir Kale and Abhinav Rastogi. 2020. Text-to-text pre-training for data-to-text tasks. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 97–102, Dublin, Ireland. Association for Computational Linguistics.

Rémi Lebret, David Grangier, and Michael Auli. 2016. Neural text generation from structured data with application to the biography domain. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1203–1213, Austin, Texas. Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Liang Li, Can Ma, Yinliang Yue, and Dayong Hu. 2021. Improving encoder by auxiliary supervision tasks for table-to-text generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5979–5989, Online. Association for Computational Linguistics.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Tianyu Liu, Kexiang Wang, Lei Sha, Baobao Chang, and Zhifang Sui. 2017. Table-to-text generation by structure-aware seq2seq learning. In *AAAI Conference on Artificial Intelligence*.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Ankur Parikh, Xuezhi Wang, Sebastian Gehrmann, Manaal Faruqui, Bhuwan Dhingra, Diyi Yang, and Dipanjan Das. 2020a. ToTTo: A controlled table-to-text generation dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1173–1186, Online. Association for Computational Linguistics.

Ankur Parikh, Xuezhi Wang, Sebastian Gehrmann, Manaal Faruqui, Bhuwan Dhingra, Diyi Yang, and Dipanjan Das. 2020b. ToTTo: A controlled table-to-text generation dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1173–1186, Online. Association for Computational Linguistics.

Ratish Puduppully, Li Dong, and Mirella Lapata. 2019. Data-to-text generation with content selection and planning. AAAI'19/IAAI'19/EAAI'19. AAAI Press.

Ratish Puduppully and Mirella Lapata. 2021. Data-to-text Generation with Macro Planning. *Transactions of the Association for Computational Linguistics*, 9:510–527.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21(1).

Clement Rebuffel, Marco Roberti, Laure Soulier, Geoffrey Scoutheeten, Rossella Cancelliere, and Patrick Gallinari. 2022. Controlling hallucinations at word level in data-to-text generation. *Data Min. Knowl. Discov.*, 36(1):318–354.

Clément Rebuffel, Laure Soulier, Geoffrey Scoutheeten, and Patrick Gallinari. 2019. A hierarchical model for data-to-text generation.

Fei Wang, Zhewei Xu, Pedro Szekely, and Muhao Chen. 2022. Robust (controlled) table-to-text generation with structure-aware equivariance learning. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5037–5048, Seattle, United States. Association for Computational Linguistics.

Sam Wiseman, Stuart Shieber, and Alexander Rush. 2017. Challenges in data-to-document generation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2253–2263, Copenhagen, Denmark. Association for Computational Linguistics.

A Appendix

A.1 Examples of Subjective Text Generation from Tables

Table:

TEAMS	M	W	L	T	PT	Series Form
Delhi Daredevils	16	11	5	0	22	WLWLW
Kolkata Knight Riders	16	10	5	0	21	WWLLW
Mumbai Indians	16	10	6	0	20	WLWWL
Chennai Super Kings	16	8	7	0	17	LWWWL
Royal Challengers Bangalore	16	8	7	0	17	LWLWW
Kings XI Punjab	16	8	8	0	16	LWLWW
Rajasthan Royals	16	7	9	0	14	LLWLW
Deccan Chargers	16	4	11	0	9	WWLLL
Pune Warriors	16	4	12	0	8	LLLLL

Reference Text:

In the league, Delhi Daredevils have showcased remarkable performance, grabbing 11 wins out of 16 matches and proving dominance with 22 points. Kolkata Knight Riders follow closely with 10 wins and 21 points. Mumbai Indians secured 10 wins as well, with a slightly fluctuating form. The competition has been intense for the teams like Chennai Super Kings and Royal Challengers Bangalore, both securing 17 points each. Kings XI Punjab, Rajasthan Royals and Deccan Chargers also fail to display promising performance with 16, 14 and 9 points each. Pune Warriors faced an uphill battle with only 8 points and securing bottom spot.

Figure 3: Generating text with subjectivity: The table shows different teams and their relative performances in a tournament. The table is described by the reference text, enriched with subjectivity. It is an example instance table from the sports domain of the **TaTS dataset**. In the table the abbreviations denote the following; M:Matches; W:Wins; L:Losses; T:Ties; PT:Points.

Figure 3 and figure 4 showcase some more examples of data instances from the **TaTS dataset**. These tables are collected from Sports and Weather forecast domains respectively.

A.2 Survey Results

To establish the social significance of this research, we employed a public survey. Students from different institutes volunteered for the survey. Also, we invited responses openly through public platforms such as LinkedIn.

The survey contained tables from four different genres such as the Sports points table (Points table), financial statement table (Pricing table), weather forecast table (Weather table), and real-estate property table (Real-estate table). Each table instance was accompanied by annotated reference text from the **TaTS dataset**. The participants were asked *Will an auto-generated text like this be helpful for you to understand the table or write something based on the table?* for each instance with the options as **Useful** (Yes), **Not sure** (Maybe), or **Not useful** (No).

Table:

Forecast Date	Valid From	Valid Till	Rainfall (mm)	TempMax (deg C)	TempMin (deg C)	Humidity-I (%)	Wind Speed (m/s)	Wind Direction (deg)	Cloud Cover (octa)
0	2024-02-07	2024-02-06	2024-02-07	0	28	20	38	10	248
1	2024-02-08	2024-02-07	2024-02-08	0	27	19	23	11	335
2	2024-02-09	2024-02-08	2024-02-09	0	26	17	17	12	9
3	2024-02-10	2024-02-09	2024-02-10	0	25	16	14	12	347
4	2024-02-11	2024-02-10	2024-02-11	0	26	16	16	11	342

Reference Text:

The data from the table shows the variation in weather forecast for consecutive 5 days in the month of February. There has been no rainfall recorded between 7th of February to 11th of February. While both the maximum and minimum degree of temperature is dropping, the region is slowly pacing towards colder days and nights. The humidity pattern demonstrates a mixed variation with a straight fall from 70% to 29% in the first three days, indicating even minor chances of rain. The weather is overall comfortable and pleasing. People in this region witness clear skies and an overall amiable weather. Significant cloud cover was observed on 11th of February which warns of the advent of scattered drizzles. The area experiences moderate breezes making the environment more conducive to the people. On a general note, the weather is fairly enjoyable, but it would be good if precautionary measures are adopted to counteract sudden showers or temperature drip.

Figure 4: Generating text with subjectivity: The table contains data on the weather forecast in the district of Kolkata from 6th February 2024 to 11th February 2024. The subjective reference text describes the table. This example data instance is taken from the **TaTS dataset**.

Approximately **2,400** participants actively engaged in the survey. Nearly **93.1%** of the survey participants disclosed association with the technology sector, through academics or profession and **74.7%** of the respondents actively acknowledged the utilization of data tables as a means to comprehend information.

Figure 1.3 depicts the results of the survey. The blue, yellow, and red refer to the Yes, Not Sure, and No responses respectively for each instance. The survey reports unequivocally underscore the profound impact and efficacy of this research in enhancing the everyday experiences of the common people.

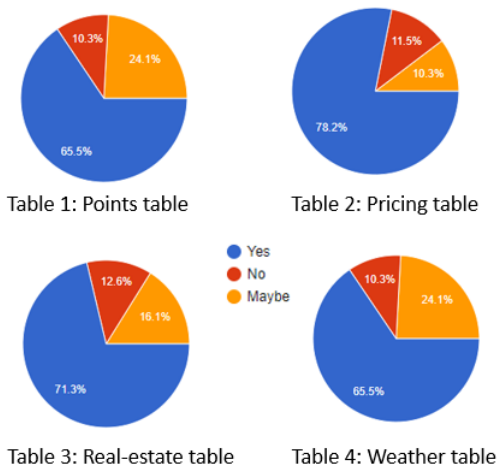

Figure 5: Survey Report: Percentage of people accepting or denying the use case of the work.

A.3 Data Collection

The tables of TaTS Dataset have been scraped from 5 sites namely, Groww [1], Indian Meteorological Dept. [2], ESPN Cric Info [3], IPL [4] and Goal [5] which are mentioned in the section 3.1.

[1] https://groww.in/
[2] https://mausam.imd.gov.in/imd_latest/contents/agromet/advisory/indiadistrictforecast.php
[3] https://www.espncricinfo.com/
[4] https://www.iplt20.com/
[5] https://www.goal.com/en-in

A.4 TaTS Dataset Statistics

The dataset contains three different kinds of tables such as financial-statement tables, weather forecast tables, and sports tables. The financial statements tables are of different companies that are listed on the National Stock Exchange, India. There are three kinds of financial tables for each company, namely its income statement, balance sheet, and cash flow tables. On the other hand, the weather forecast data is collected from the official site of

Govt. of India. These tables present data for each district in each state. And the tables of football and cricket games from different tournaments. Statistics of the dataset have been shown in table 3. As mentioned in section 3.1, a total of 3077 tables are collected with an average of nearly 9 rows and nearly 7 features.

A.5 Dataset Annotation

The data instances have been manually annotated which is mentioned in section 3.2. The annotators were given a detailed instruction set to follow and were paid a total of 1.20$ per data instance for annotation. The annotation task was conducted through regular monitoring over two months, following a fixed set of instructions. We asked the annotators to write at most 150 words for each instance. We try to keep the length of the reference text invariable for a variable number of rows and columns.

We define the annotation task in three steps: (1) Information extraction from the table, (2) Subjectivity infusion, and (3) Text generation. In step 1, the annotators are expected to extract the objective information in the table and understand its dynamics, changes, and relations. Furthermore, they should understand the emotions lying with the variables in the data and infuse subjectivity into the information. In the last step, they should write a text combined with the previous semantics which describes the table.

A.6 Table Linearization

Linearization of tables is crucial when it is given as input in seq-to-seq models. We used a tagged concatenation of delimiter separated row-major and column-major ordering to generate a flattened sequence of table. The row-major and column-major orders are initiated with tags **<rows>** and **<columns>**. Whereas, each row in row-major order and each column in column-major order is separated by **<SEP>** token. The intuition behind the linearization is described in section 4.1.

A row-major flattening of the table in figure 6 can be written as, *Nottinghamshire : Matches 10 | Wins 7 | Losses 1 ... <SEP> Lancashire : Matches 10 | Wins 5 | ... <SEP>*.

Similarly, a column-major order is also generated. The row-major and column-major flattenings are concatenated with a delimiter token. **<rows>** and **<columns>** tags are used to separate both orders. After trying other linearization techniques as

	0	1	2	3	4	6	7	8
0	TEAMS	M	W	L	T	PT	NRR	Series Form
1	Nottinghamshire	10	7	1	0	16	1.315	WLWWW
2	Lancashire	10	5	3	0	12	-0.250	LLWWL
3	Leicestershire	10	4	3	0	11	-0.180	WWLLW
4	Durham	10	4	5	0	9	0.421	LWWWW
5	Yorkshire	10	3	5	0	8	0.297	WLLLL
6	Derbyshire	10	1	7	0	4	-1.583	LLWLL

Figure 6: Table flattening: extracting information from row-major and column-major ordering

well, we choose this approach due to better consistency in generated texts.

A.7 Experimental Setup

The T5-small, base, and large models are of approximately 60 million, 220 million, and 770 million parameters respectively. We used the Nvidia RTX A6000 GPU to train the models where each epoch took nearly 15 minutes to complete training and validation testing.

A.8 Qualitative Analysis

Qualitative analysis was performed on the generated samples. The evaluation guideline provided to the annotators included three parameters, **Coherence**, **Coverage and Accuracy**, and **Subjectivity Capture**. They were asked to mark the generations from 1 to 10, 1 being the lowest in quality and 10 being the highest in quality. Subjectivity capture measurement refers to the correct emotion being captured in the subjective phrases present in the reference text. The reference text may not introduce emotion regarding every objective information. However, it is not penalized for that. We only take into account the correctness of the generated phrases. Coverage and Accuracy is measured as the intersection of information present in the table and reflected in the reference text. The higher the intersection, the greater the score. Coherence is evaluated by the readers' comprehension of the text and the probability of each statement being inferred from the previous ones.

The text in figure 7 was generated from the table shown above by the T5-large model. The phrases marked with green show the subjectivity that is captured fruitfully at the right places. However, some wrong objective information was also generated which is marked in red. It signifies that the model can understand the semantic meaning of the data

Genre	No. of tables	Avg. no. of rows	Avg. no. of columns
Finance			
Income statement	500	11	5
Balance Sheet	500	11	5
Cash flow	500	11	5
Weather forecast			
District-wise tables	1168	5	10
Sports			
Big Bash	12	8	7
IPL	16	8.5	8
CPL	11	6.36	6
Vitality Blast	53	7.16	6
ICC World Cup	27	5.667	6
ICC Women's World Cup	17	5.882	6
PSL	8	5.75	6
Women's Big Bash	8	8	6
Other Cricket Tournaments	33	8.969	6.272
Football Tournaments	224	11.18	8
Total	**3077**		

Table 3: Sources of different web-scraped tables with the average number of rows and columns.

present in the table. See section 5.2 for details.

	0	1	2	3	4	6	7	8
0	TEAMS	M	W	L	T	PT	NRR	Series Form
1	Somerset	14	12	2	0	24	1.460	WWWWW
2	Hampshire	14	9	5	0	18	0.820	WWWLW
3	Surrey	14	8	6	0	16	1.192	LLLWL
4	Essex	14	8	6	0	16	0.088	WLLLL
5	Kent	14	7	7	0	14	0.287	LLWWW
6	Sussex	14	6	8	0	12	-0.871	LWWWL
7	Gloucestershire	14	5	9	0	10	-0.993	LWLLW
8	Glamorgan	14	5	9	0	10	-1.060	LLLLL
9	Middlesex	14	3	11	0	6	-0.932	WWLWL

In this competitive season, Somerset has taken the lead with 21 points from 14 matches, showcasing their dominance with 10 wins and a net run rate of 0.045. Their win streak underscores their dominant form. Essex closely follows with 18 points, winning 10 matches and an identical number of wins. Sussex and Surrey are also in the mix, with 15 and 12 points each. Gloucestershire, Glamorgan, Sussex, and Kent are facing challenges but determined to improve their performance. As the tournament progresses, teams battle for the top spotless competition. The competition remains fierce, promising more exciting encounters as the season unfolds. For an eventful season promises more thrilling action ahead.

Figure 7: Qualitative Analysis: The text is generated by the model from the given table. The green phrases mark the capture of subjectivity occurring in the right places. Red phrases mark misinformation generated.