Table to Text Generation with Subjectivity and Objectivity

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**. A new linearization technique is implemented to flatten the tables. We perform the task using various sequence-to-sequence models 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

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

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 1: Generating text with subjectivity: The points table is described by the text shown below where the green marked phrases refer to the infused subjectivity. These phrases come from the comprehension of the table data. (Due to space constraints, shown in figure 4)

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.

The table depicted in Figure 1 presents a tournament points table with 9 teams with the number of matches, wins, losses, ties, points, and series forms respectively. Within the accompanying narrative, a number of expressions are employed, which does 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 ***remarkable performance*** comes from the idea that winning 11 matches out of 16 is **good**. Whereas, ***closely*** comes from the understanding that being one point behind the table topper is not a huge difference. On the other hand, ***uphill battle*** is a phrase that can be used from the idea that winning just 4 matches out of 16 is **bad**. Similarly example can be seen in section A (figure 5).

So, the goal of the task is to generate text that evidently describes the objective information but with a touch of subjectivity or interpretation that is not straightforward from the table data.

1.2 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. Nearly **93.1%** of the survey participants disclosed association with the technology sector and **74.7%** of the respondents actively acknowledged the routine utilization of data tables as a means to comprehend information.

In light of the statistics from figure 6 (see A.2), the survey unequivocally underscores the profound impact and efficacy of this research in enhancing the everyday experiences of the common people.

1.3 Our Contributions

Our contributions are:

1. The formulation of an entirely novel problem statement of subjective text generation from tables, enriched with social significance.
2. The creation of a limited-sample **TaTS** dataset containing 409 instances of tables from different sports genres and with different features.
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 a number of 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) 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.

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

We collected the tables using various web-based sports sites that are entitled to showcase different game scores. Three sites were selected as sources of the tables such as ESPN Cric Info, IPL, and Goal (see A.3 for sources). The tables are scraped using Python Programming language.

A total of 328 tables were collected from the sources by web scraping. These tables refer to the points table of 89 different tournaments. There are a total of 143 football points table collected and 185 cricket points table. The complete TaTS dataset statistics can be found in the table 3 in the Appendix section (see A.4).

3.2 Data Filtration

The collected data is filtered manually. Tables having less than 4 teams or less than 3 features were removed. Also if 25% of the teams have the highest and lowest score simultaneously or any team has played less than 2 matches, those were also removed. After the data filtering, the number of tables in the dataset is reduced from 328 to 301.

3.3 Synthetic Generation

Due to the scarcity of data instances, a few samples were generated synthetically by ChatGPT. Various different zero-shot prompts were tried to get an accurate result. The most useful prompt was *Generate a tournament point table with arbitrary team names and columns as Matches, win, loss, draw, points, series form having more than 6 teams, more than 10 matches*. The number of teams and number of matches in the prompt was varied to keep symmetry in the dataset. An example of such a table is shown in figure 7 (See A.5). A total of 108 tables were created by ChatGPT which were added to the previous dataset. So the number of instances in the complete **TaTS dataset** is 409 (see table 4).

3.4 Annotation Details

We employ three annotators for this task who are proficient in the English language. One of the annotators is a graduate of Computer science and engineering while the others are PhD students and postgraduates in English literature respectively.

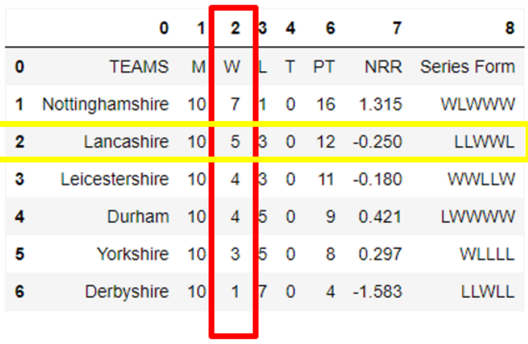
They were provided with detailed annotation guidelines with rules, objectives, and multiple examples. Each instance in the set was annotated only once. The distribution among the annotators was 100, 201, and 108 instances. An example annotation is shown in figure 8 (See A.6). 30 samples were annotated by all 3 annotators which are used to compute the Bleu-2 score for **Inter-Annotator Agreement** (IAA). The pair-wise average B-2 score comes as **37.28** (A-B), **48.29** (B-C), and **32.59** (A-C). 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 perform the task using sequence-to-sequence models. Due to the relatively small dataset size, 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, the performance of an individual team and its relative standing within the table against other teams. Hence, the application of both row-major flattening and column-major flattening techniques is required.



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

A row-major flattening of the table 2 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. The table in figure 2 can be flattened in a sequence as shown in figure 9 (See A.7).

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 3.

4.3 Experimental Setup

The T5 models are fine-tuned with the prefix *Describe the table briefly:* for 50 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.8). A train-valid-test split of 80%-10%-10% is used.

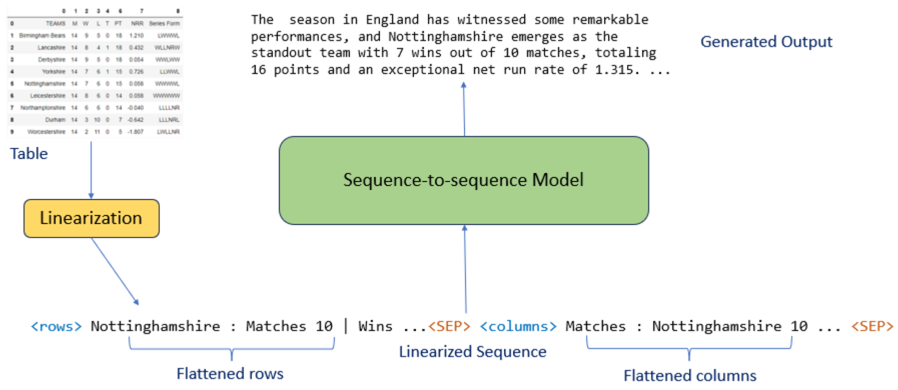


Figure 3: Architecture: Tables are flattened through linearization which is then appended with delimiter tokens. This sequence acts as input to various sequence-to-sequence models which generates the expected output.

5 Results

We conducted experiments with various sequence-to-sequence models and compared the results.

5.1 Quantitative Analysis

Model	Bleu-4	Meteor	Rouge-L
T5-small	11.82	14.40	7.36
T5-base	13.17	18.51	10.08
T5-large	17.12	19.02	11.24
BART	10.42	12.33	9.54

Table 1: 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.

The best Bleu-4 metric score achieved is around **17.12**. 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.

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 10 shows an example of generated output by the model (see in section A.9). 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.)*. They were asked to mark the generations out of 10 (10 being highest and 1 being lowest) for each feature and the average scores are shown in table 2.

	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 clearly understandable the generations are able to 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. To facilitate our goal, we curate the **TaTS** Dataset, develop a new 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, it is evident that there is ample room for further exploration. Expanding the task to encompass a larger and more diverse dataset, comprising tables from various 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 a single genre. Subjectivity in the same type of table can be influenced by similar semantics, which may not be the case for a mixed dataset. Hence, the performance may vary based on the genre of tables. Moreover, we developed a new table flattening technique to fit tables in sequence-to-sequence models which are trained on normal text. For even larger tables, a table-specific architecture may be needed.

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.

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

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

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 4: Generating text with subjectivity: The points table is described by the text shown below where the green marked phrases refer to the infused subjectivity. These phrases come from the comprehension of the table data.

The figure 4 is an extended image of the figure 1 shown in section 1.1 for better visibility.

Stock Name	Price (USD)	Traded Amount (Shares)	Change from Last Week (%)
Apple Inc.	150.00	2,500,000	+2.5%
Microsoft	280.50	1,800,000	-1.2%
Amazon.com	3450.00	1,200,000	+3.8%
Google (Alphabet)	2600.00	1,500,000	+0.5%
Tesla Inc.	720.50	1,750,000	-0.8%
Facebook	330.20	1,600,000	+1.9%
Netflix	500.75	1,100,000	-0.3%
NVIDIA Corp.	220.80	1,300,000	+4.1%
Johnson & Johnson	170.40	900,000	+1.2%
Pfizer Inc.	45.60	1,000,000	+0.7%

The stock market reveals a dynamic landscape with notable changes this week. Apple Inc. exhibits a positive trend with a 2.5% increase in its stock price, while Microsoft sees a slight dip of -1.2%. Amazon.com surges with an impressive +3.8% change, reflecting investor optimism, and Google (Alphabet) maintains a steady climb with a +0.5% increase. Tesla Inc. experiences a modest decrease of -0.8%, while Facebook gains +1.9% in its stock value. Netflix observes a minor -0.3% change, while NVIDIA Corp. soars with a significant +4.1% rise. Johnson & Johnson and Pfizer Inc. both demonstrate positive movements, with increases of +1.2% and +0.7%, respectively. These fluctuations in stock prices evoke a range of reactions among investors, from excitement to caution, as they navigate the ever-changing financial landscape.

Figure 5: Generating text with subjectivity: The price change table is described in the reference text below which contains subjectivity.

Figure 5 shows another example of the problem statement. The table comprises changes in prices with a subjective reference text describing the table (manually annotated). This table signifies the utility of this work in other genres as well.

A.2 Survey Results

To establish the social significance of this research, we conducted a survey with four example tables from different genres and reference texts respectively. The four genres are tournament points table, price change table, real-estate property details table, and weather report table. The survey participants were asked to mark each instance as **Useful** (Yes), **Not sure** (Maybe), or **Not useful** (No) whose results are shown in figure 6. To see other survey details, please refer to section 1.2.

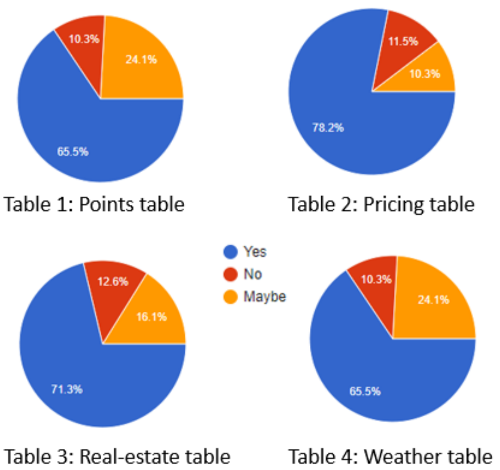


Figure 6: Survey Report: Percentage of people accepting or denying the use case of the work.

A.3 Data Collection

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

A.4 TaTS Dataset Statistics

The dataset contains 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 328 tables are collected with an average of nearly 9 teams having nearly 7 features.

[1] https://www.espncricinfo.com/
[2] https://www.iplt20.com/
[3] https://www.goal.com/en-in

Tournament	No. of tables	Avg. no. of rows	Avg. no. of columns
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	143	11.18	8
Total	328	8.95	7.03

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

A.5 Complete TaTS Dataset with Synthetically Generation

A number of tables in the TaTS dataset are generated synthetically as mentioned in the section 3.3. An example of such a synthetically generated table is shown in figure 7. The table contains 6 teams with 6 features.

Teams	Matches	Wins	Losses	Draws	Points	Series Form
Team Alpha	12	8	3	1	25	WWLWLWWDWWLW
Team Bravo	11	6	3	2	20	WWLWDDWLWDL
Team Charlie	10	4	5	1	13	LLWLWDWLDD
Team Delta	12	10	2	0	30	WWWWWWWWWWWL
Team Echo	11	3	7	1	10	LWLWLDLLDLL
Team Foxtrot	13	9	3	1	28	WWLWLWWWLWWLL

Figure 7: Example of a synthetically generated table

Having generated 108 table instances synthetically, the size of the complete TaTS dataset stands at **409** which is shown in table 4.

	Number of Instances
In scraped dataset	301
Generated by ChatGPT	108
In complete dataset	409

Table 4: Complete dataset statistics with web scraped tables and synthetically generated tables

A.6 Manual Annotation Example

The data instances have been manually annotated which is mentioned in section 3.4. The annotators were given a detailed instruction set to follow and were paid a total of 0.20$ per data instance for annotation. An example of such an annotation is given in figure 8.

	0	1	2	3	4	6	7	8
0	TEAMS	M	W	L	T	PT	NRR	Series Form
1	Birmingham Bears	14	9	5	0	18	1.210	LWWWL
2	Lancashire	14	8	4	1	18	0.432	WLLNRW
3	Derbyshire	14	9	5	0	18	0.054	WWLWW
4	Yorkshire	14	7	6	1	15	0.726	LLWWL
5	Nottinghamshire	14	7	6	0	15	0.058	WWWWL
6	Leicestershire	14	8	6	0	14	0.058	WWWWW
7	Northamptonshire	14	6	6	0	14	-0.040	LLLLNR
8	Durham	14	3	10	0	7	-0.642	LLLNRL
9	Worcestershire	14	2	11	0	5	-1.807	LWLLNR

The season has been a rollercoaster ride, with some impressive team performances. Birmingham Bears lead the pack, securing 9 wins out of 14 matches, earning 18 points, and boasting an outstanding net run rate of 1.210. Lancashire and Derbyshire are hot on their heels, also tallying 18 points each. Lancashire with 8 wins and Derbyshire with 9 wins showcased their mettle. Yorkshire and Nottinghamshire both ended with 15 points, displaying some gripping battles. Leicestershire and Northamptonshire are tied at 14 points, with Leicestershire enjoying the advantage of a better net run rate. Durham, Worcestershire, had their challenges but put up a spirited fight throughout the series.

Figure 8: Example of table and reference text annotation

A.7 Example of Table Linearization

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 process of linearization is described in section 4.1.

Here in figure 9 an example of linearization of the table in figure 2 is shown.

A.8 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.9 Qualitative Analysis

Qualitative analysis was performed on the generated samples. The text in figure 10 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 present in the table. See section 5.2 for details.

```
<rows> Nottinghamshire : Matches 10 | Wins 7 | Losses 1 |
Ties 0 | Points 16 | NRR 1.315 | Series Form WLWWW <SEP>
Lancashire : Matches 10 | Wins 5 | Losses 3 | Ties 0 |
Points 12 | NRR -0.25 | Series Form LLWWL <SEP>
Leicestershire : Matches 10 | Wins 4 | Losses 3 | Ties 0 |
Points 11 | NRR -0.18 | Series Form WWLLW <SEP> Durham :
Matches 10 | Wins 4 | Losses 5 | Ties 0 | Points 9 | NRR
0.421 | Series Form LWWWW <SEP> Yorkshire : Matches 10 |
Wins 3 | Losses 5 | Ties 0 | Points 8 | NRR 0.297 | Series
Form WLLLL <SEP> Derbyshire : Matches 10 | Wins 1 | Losses
7 | Ties 0 | Points 4 | NRR -1.583 | Series Form LLWLL
<SEP>  <columns> Matches : Nottinghamshire 10 | Lancashire
10 | Leicestershire 10 | Durham 10 | Yorkshire 10 |
Derbyshire 10 <SEP> Wins : Nottinghamshire 7 | Lancashire 5
| Leicestershire 4 | Durham 4 | Yorkshire 3 | Derbyshire 1
<SEP> Losses : Nottinghamshire 1 | Lancashire 3 |
Leicestershire 3 | Durham 5 | Yorkshire 5 | Derbyshire 7
<SEP> Ties : Nottinghamshire 0 | Lancashire 0 |
Leicestershire 0 | Durham 0 | Yorkshire 0 | Derbyshire 0
<SEP> Points : Nottinghamshire 16 | Lancashire 12 |
Leicestershire 11 | Durham 9 | Yorkshire 8 | Derbyshire 4
<SEP> NRR : Nottinghamshire 1.315 | Lancashire -0.25 |
Leicestershire -0.18 | Durham 0.421 | Yorkshire 0.297 |
Derbyshire -1.583 <SEP> Series Form : Nottinghamshire WLWWW
| Lancashire LLWWL | Leicestershire WWLLW | Durham LWWWW |
Yorkshire WLLLL | Derbyshire LLWLL <SEP>
```

Figure 9: Example of flattened table (See table in figure 2)

	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 10: 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.