Deep Learning for Financial Question Answering

Max Kan and Taichi Kato

Introduction: Problem

Financial documents are high-volume, complicated, with tables and figures

Current challenge: extracting **specific insights** from dense information



FORM 10-Q: What is the year-on-year revenue growth in Tesla's automotive sales? (%)

Tesla, Inc. Notes to Consolidated Financial Statements (unaudited)

Note 1 - Summary of Significant Accounting Policies

Unaudited Interim Financial Statements

The consolidated financial statements of Tesla, Inc. ("Tesla", the "Company", "we", "us" or "our"), including the consolidated balance sheet as of March 31, 2023, the consolidated statements of operations, the consolidated statements of comprehensive income, the consolidated statements of redeemable noncontrolling interests and equity, and the consolidated statements of cash flows for the three months ended March 31, 2023 and 2022, as well as other information disclosed in the accompanying notes, are unaudited. The consolidated balance sheet as of December 31, 2022 was derived from the audited consolidated financial statements and the accompanying notes should be read in conjunction with the annual consolidated financial statements and the accompanying notes contained in our Annual Report on Form 10-K for the year ended December 31, 2022.

The interim consolidated financial statements and the accompanying notes have been prepared on the same basis as the annual consolidated financial statements and, in the opinion of management, reflect all adjustments, which include only normal recurring adjustments, necessary for a fair statement of the results of operations for the periods presented. The consolidated results of operations for any interim period are not necessarily indicative of the results to be expected for the full year or for any other future years or interim periods.

Reclassifications

Certain prior period balances have been reclassified to conform to the current period presentation in the accompanying notes.

Revenue Recognition

Revenue by source

The following table disaggregates our revenue by major source (in millions):

		Three Months Ended March 31,			
	2023	2023		2022	
Automotive sales	\$	18,878	S	15,514	
Automotive regulatory credits		521		679	
Energy generation and storage sales		1,413		503	
Services and other		1,837		1,279	
Total revenues from sales and services		22,649		17,975	
Automotive leasing		564		668	
Energy generation and storage leasing		116		113	
Total revenues	\$	23,329	\$	18,756	

Automotive Segment

Automotive Sales Revenue

The total sales return reserve on vehicles sold with resale value guarantees was \$68 million and \$91 million as of March 31, 2023 and December 31, 2022, respectively, of which \$34 million and \$40 million was short-term, respectively,

Deferred revenue is related to the access to our Full Self Driving ("FSD") features and ongoing maintenance, internet connectivity, free Supercharging programs and over-the-air software updates primarily on automotive sales, which amounted to \$3.04 billion and \$2.91 billion as of March 31, 2023 and December 31, 2022, respectively.

Deferred revenue is equivalent to the total transaction price allocated to the performance obligations that are unsatisfied, or partially unsatisfied, as of the balance sheet date. Revenue recognized from the deferred revenue balance as of December 31, 2022 and 2021 was \$134 million and \$66 million for three months ended March 31, 2023 and 2022, respectively. Of the total deferred revenue balance as of March 31, 2023, we expect to recognize \$679 million of revenue in the next 12 months. The remaining balance will be recognized at the time of transfer of control of the product or over the performance period.

FORM 10-Q: What is the year-on-year revenue growth in Tesla's automotive sales? (%)

Tesla, Inc. Notes to Consolidated Financial Statements (unaudited)

Note 1 - Summary of Significant Accounting Policies

Unaudited Interim Financial Statements

The consolidated financial statements of Tesla, Inc. ("Tesla", the "Company") we", "us" or "our"), including the consolidated balance sheet as of March 31, 2023, the consolidated statements of operations, the consolidated statements of comprehensive income, the consolidated statements of redeemable noncontrolling interests and equity, and the consolidated statements of cash flows for the three months ended March 31, 2023 and 2022, as well as other information disclosed in the accompanying notes, are unaudited. The consolidated balance sheet as of December 31, 2022 was derived from the audited consolidated financial statements and the accompanying notes should be read in conjunction with the annual consolidated financial statements and the accompanying notes on tail and the accompanying notes of the statements are statements and the accompanying notes of the statements are statements as of the statements and the accompanying notes of the statements are statements and the accompanying notes of the statements are statements are statements and the accompanying notes of the statements are statements and the accompanying notes of the statements are statements are statements and the accompanying notes of the statements of t

The interim consolidated financial statements and the accompanying notes have been prepared on the same basis as the annual consolidated financial statements and, in the opinion of management, reflect all adjustments, which include only normal recurring adjustments, necessary for a fair statement of the results of operations for the periods presented. The consolidated results of operations for any interim period are not necessarily indicative of the results to be expected for the full year or for any other future years or interim periods.

Reclassifications

Certain prior period balances have been reclassified to conform to the current period presentation in the accompanying notes.

Revenue Recognition

Revenue by source

The following table disaggregates our revenue by major source (in millions)

	Three Months Ended March 31,				
		2023		2022	
Automotive sales	\$	(18,878)	S	(15,514)	
Automotive regulatory credits		521		679	
Energy generation and storage sales		1,413		503	
Services and other		1,837		1,279	
Total revenues from sales and services		22,649		17,975	
Automotive leasing		564		668	
Energy generation and storage leasing		116		113	
Total revenues	\$	23,329	\$	18,756	

Automotive Segment

Automotive Sales Revenue

The total sales return reserve on vehicles sold with resale value guarantees was \$68 million and \$91 million as of March 31, 2023 and December 31, 2022, respectively, of which \$34 million and \$40 million was short-term, respectively.

Deferred revenue is related to the access to our Full Self Driving ("FSD") features and ongoing maintenance, internet connectivity, free Supercharging programs and over-the-air software updates primarily on automotive sales, which amounted to \$3.04 billion and \$2.91 billion as of March 31, 2023 and December 31, 2022, respectively.

Deferred revenue is equivalent to the total transaction price allocated to the performance obligations that are unsatisfied, or partially unsatisfied, as of the balance sheet date. Revenue recognized from the deferred revenue balance as of December 31, 2022 and 2021 was \$134 million and \$66 million for three months ended March 31, 2023 and 2022, respectively. Of the total deferred revenue balance as of March 31, 2023, we expect to recognize \$679 million of revenue in the next 12 months. The remaining balance will be recognized at the time of transfer of control of the product or over the performance period.

FORM 10-Q: What is the year-on-year revenue growth in Tesla's automotive sales? (%)

Tesla, Inc. Notes to Consolidated Financial Statements (unaudited)

Note 1 - Summary of Significant Accounting Policies

Unaudited Interim Financial Statements

The consolidated financial statements of Tesla, Inc. ("Tesla", the "Company") we", "us" or "our"), including the consolidated balance sheet as of March 31, 2023, the consolidated statements of operations, the consolidated statements of comprehensive income, the consolidated statements of redeemable noncontrolling interests and equity, and the consolidated statements of cash flows for the three months ended March 31, 2023 and 2022, as well as other information disclosed in the accompanying notes, are unaudited. The consolidated balance sheet as of December 31, 2022 was derived from the audited consolidated financial statements as of that date. The interim consolidated financial statements and the accompanying notes should be read in conjunction with the annual consolidated financial statements and the accompanying notes contained in our Annual Report on Form 10-K for the year ended December 31, 2022.

The interim consolidated financial statements and the accompanying notes have been prepared on the same basis as the annual consolidated financial statements and, in the opinion of management, reflect all adjustments, which include only normal recurring adjustments, necessary for a fair statement of the results of operations for the periods presented. The consolidated results of operations for any interim period are not necessarily indicative of the results to be expected for the full year or for any other future years or interim periods.

Reclassifications

Certain prior period balances have been reclassified to conform to the current period presentation in the accompanying notes.

Revenue Recognition

Revenue by source

The following table disaggregates our revenue by major source (in millions

	Three Months Ended March 31,			
	2023	2022	_	
Automotive sales	\$ 18,878	\$ (15,51	4	
Automotive regulatory credits	521	67	مرو	
Energy generation and storage sales	1,413	50.	3	
Services and other	1,837	1,27	9	
Total revenues from sales and services	22,649	17,97	5	
Automotive leasing	564	66	8	
Energy generation and storage leasing	116	11:	3	
Total revenues	\$ 23,329	\$ 18,75	6	

Automotive Segment

8878 / 15514 = 121% cember 31, 2022, respectively, of which \$34 million and \$40 million was short-term, respectively upercharging programs and over-the-air software updates primarily on automotive sales, which

cember 31, 2022, respectively, of which \$34 million and \$40 million was short-term, respectively.

Deferred revenue is equivalent to the total transaction price allocated to the performance obligations that are unsatisfied, or partially unsatisfied, as of the balance sheet date. Revenue recognized from the deferred revenue balance as of December 31, 2022 and 2021 was \$134 million and \$66 million for three months ended March 31, 2023, we expect to recognize \$679 million of revenue in the next 12 months. The remaining balance will be recognized at the time of transfer of control of the product or over the performance period.

Previous Solutions

Extractive QA

- SQuAD Dataset: A generic QA dataset with question-answer pairs derived from Wikipedia articles
- Transformers: T5/Roberta, etc.
- Symbolic operators: Mathematica,
 Wolframalpha, etc.
- Limitations: Limited to general QA, not specific to **financial data**, structured to be extractive and not generative

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called "showers".

What causes precipitation to fall? gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail? graupel

Where do water droplets collide with ice crystals to form precipitation?

within a cloud

Our Solution — Financial Numerical Reasoning Dataset

FINQA: A Dataset of Numerical Reasoning over Financial Data

Zhiyu Chen¹, Wenhu Chen¹, Charese Smiley², Sameena Shah²,
Iana Borova¹, Dylan Langdon¹, Reema Moussa¹, Matt Beane¹, Ting-Hao Huang³,
Bryan Routledge⁴ and William Yang Wang¹
¹University of California, Santa Barbara
²J.P. Morgan
³Pennsylvania State University
⁴Carnegie Mellon University
{zhiyuchen, william}@cs.ucsb.edu

Abstract

The sheer volume of financial statements makes it difficult for humans to access and analyze a business's financials. Robust numerical reasoning likewise faces unique challenges in this domain. In this work, we focus on answering deep questions over financial data, aiming to automate the analysis of a large corpus of financial documents. In contrast to existing tasks on general domain, the finance domain includes complex numerical reasoning and understanding of heterogeneous representations. To facilitate analytical progress, we propose a new large-scale dataset, FINQA, with Question-Answering pairs over Financial

by an exponentially expanding collection of company financial documents (MacKenzie et al., 2012; Lange et al., 2016) such that it is genuinely unclear whether dedicated human effort can produce fiscal analysis of sufficient quality for current decision making. This poses an interesting question: can we automate such deep analysis of financial data?

A few NLP studies in Question Answering (QA) explored the numerical reasoning capabilities needed to answer questions correctly. For example, the DROP dataset (Dua et al., 2019) focused on Wikipedia-based questions that require numerical reasoning, e.g., "Where did Charles travel to

Dataset: Modified <u>FinQA</u> - comprehensive for financial Q&A (Chen et. al 2021)

Contains earnings reports of S&P 500 companies, 8,281 Q&A pairs

Numerical reasoning processes included

Our Solution — Generative QA

For our task, we need to generate/synthesize a solution based on the given context, but we may not be able to find an exact solution in the context. So our task is inherently a generative task.

We can define the task as modelling the mapping between context/question pair with an answer.

The most common model for these tasks is generally known as causal language models, and the most popular example of it is GPT.

Data Pre-processing

Context

Netor to Co	nolidated Financial Statements				
	(unaudited)				
Note 1 - Summary of Significant Accounting Policies					
Cleanding Steerin Financial Statements					
The consolidated financial statements of Torks, bo. ("Torks", the "Company", "we", "we" or "our"), and competituate incomes, the consolidated statements of redormable necessariling intercess and quipt, and the consociation of	olidated instructure of each flows for the three mon malities commissioned theoretic interprets in of their	he ended March 31, 29 date. The interior com-	123 and 2022, as we obtained financial of	fil as other info	enution disclosed in the
The interior connolidated financial statements and the accompanying notes have been proposed on the ca- normal recurring adjunctures, occurring for a fair statement of the coults of specialists for the potoda presented. T for any other financial recurrence to election portods.	ne basis as the annual consolidated financial status to consolidated souths of operations for any interior	ents and, in the opinion period are not necess	n of management, o city indicative of the	offset all adjust se nosalts to be-	terests, which include or expected for the full year
Enclaratington					
Certain prior period balances have been reclassified to conform to the current period presentation in the					
COMPANIES AND	According to the				
Revenue Recognition					
Research source					
The following table disaggregates our revenue by major source (in milliona):					
				Ended March 16,	2962
Automotive sales			19,979		15,514
Automotive regulatory credits			171		629
Energy conception and storage sales			1.40		587
Services and other			1,832		1.229
Total progress from sales and services			22,649		17,925
Automotive leasing			564		660
Enemy generation and storage leaving			116		112
Total revenues		8	23,329	8	18,156
Automotive Segment					
Automative Sales Revenue					
The total sales return reserve on vehicles sold with resale value guarantees was \$68 million and \$91 mil	ion as of March 31, 2023 and Documber 31, 2022,	espectively, of which:	SH million and \$40	million was of	lonters, respectively.
Defined revenue is related to the access to our Full Self Deiving ("PSD") Statures and ongoing mainten amounted to \$3.04 billion and \$2.01 billion as of March 31, 2023 and December 31, 2022, respectively.	nco, internet connectivity, free Supercharging proj	name and ever the air	ochrane updatos pr	inarity on auto	motive sales, which
Deferred revenue is equivalent to the total transaction price allocated to the performance obligations the December 31, 2022 and 2021 was \$154 million and \$66 million for three recents ended March \$1,2023 and 2022. Moreover, the price of the product or work the product or work the price of the product or work the profession.	expectively. Of the total deferred revenue balance a	lance sheet date. Reve s of March 31, 2025, v	nae recognised from re-expect to recogn	n the deferred to the \$679 million	revenue halance as of n of revenue in the next I

Question

what was the the interest expense in 2009?

Answer

380 million

Numerical reasoning

3.8m / 1% = 380m

Data Pre-processing

Context

	(unaudited)					
Note 1 - Summary of Significant Accounting Policies						
Consultinal Seturitor Financial Statements						
The consolidated financial statements of Torks, Inc. ("Torks", the "Company", "we", "ne" or composituative income, the consolidated statements of redomable necessariling interests and equity, as accompanying notes, are associated. The consolidated behaves show as of December 31, 2022 was derive should be used in conjunction with the associal consolidated financial statements and the accompanying as	of the consolidated matements of each flows for the ti-	have months ended March 7 as of that date. The interior	1, 2823 and 2022, as we	dl as other informeti	ion disclosed in the	
The interior consolidated financial statements and the accompanying notes have been proport normal recurring adjustments, receiving for a fair statement of the results of operations for the periods per for any other finance years or intents periods.	d on the same basis as the annual consolidated finance counted. The consolidated results of operations for a	cial statuments and, in the op my interim period are not no	einion of management, o constrily indicative of the	offset all adjustment ac results to be exper	a, which include on dod for the full year	
Resignatifications						
Certain prior period balances have been soclassified to conform to the current period presents	tion in the accompanying notes.					
Errora Econositos						
Account to source						
The following table disappropriates our propriate by major passon for millional:						
The increased ment configuration continues by trades possess by managed						
			Three Months Ended March 16.			
			303		2002	
Automotive sales			18,879	- 5	15,514	
Automative regulatory credits			521		679	
Energy generation and storage sales			1,413		503	
Services and other			1,837		1,279	
Total revenues from sales and services			22,649		17,975	
Automative leasing			564		660	
Energy generation and storage lessing			116		113	
Total revenues		8	23,329	8	18,156	
Assemble Sumor						
Automative Sales Environe						
The total sales return reserve on vehicles sold with resale value guarantees was \$68 million as	nd S91 million as of March 31, 2023 and Documber 3	is, 2022, respectively, of wi	rick \$34 million and \$40	d million was short t	arm, respectively.	
Deferred revenue is related to the access to our Full Solf Driving (*PSD*) features and ongoin amounted to \$3.04 billion and \$2.54 billion as of March 31, 2023 and Documber 31, 2022, respectively.	ng maintonance, internet connectivity, floo Superchar	ging programs and ever the	air softmans updates pr	inarity on automatic	so sales, which	
Defende revenue is equivalent to the total transaction price allocated to the performance oblig December 31, 2022 and 2021 was \$154 million and \$66 million for three meets coded March 31, 2025	and 2822, respectively. Of the total deferred revenue	of the halance sheet date. I balance as of March 31, 20	Sevenue recognised fire 25, we expect to recogn	n the deferred reven are \$679 million of a	so halance as of revenue in the next I	

Ouestion

what was the the interest expense in 2009?

Answer

380 million

Numerical reasoning

3.8m / 1% = 380m

```
{"prompt":"Interest rate to a variable interest
rate based on the three-month libor plus 2.05% (
2.05 % ) ( 2.34% ( 2.34 % ) as of october 31 ,
2009 ) .\nif libor changes by 100 basis points
our annual interest expense would change by $ 3.8
million .\nforeign currency exposure as more fully
described in note 2i .\nin the notes to
consolidated financial statements contained in
item 8 of this annual report on form 10-k , we
regularly hedge our non-u.s .\ndollar-based
exposures by entering into forward foreign
currency exchange contracts .\nthe terms of these
contracts are for periods matching the duration of
"$ 6427\", \"$ -23158 ( 23158 )\"], [\"fair value
of forward exchange contracts after a 10% ( 10 % )
unfavorable movement in foreign currency exchange
rates asset ( liability )\", \"$ 20132\", \"$
-9457 ( 9457 )\"], [\"fair value of forward
```

exchange contracts after a 10% (10 %) favorable movement in foreign currency exchange rates liability\", \"\$ -6781 (6781)\", \"\$ -38294 (38294)\"]]\n

```
q: what was the the interest expense in 2009?,
```

"completion":" a: 380 million"}

Fine-tuning with pre-trained LLM

ModFinQA PROMPTS

"prompt":"Interest rate to a variable interest rate based on the three-month libor plus 2.05% (2.05% () (2.05% () (2.03%) () (2.34%) () as of october 31, 2.09)), wiff libor changes by 100 basis points, our annual interest expense would change by \$ 3.8 million \,\nforeign currency exposure as more fully described in note 2i \,\nin the notes to consolidated financial statements contained in item 8 of this annual report on form 10-k, we regularly hedge our non-u.s \,\ndollar-based exposures by entering into forward foreign currency exchange contracts \.\nthe terms of these contracts are for periods matching the duration of

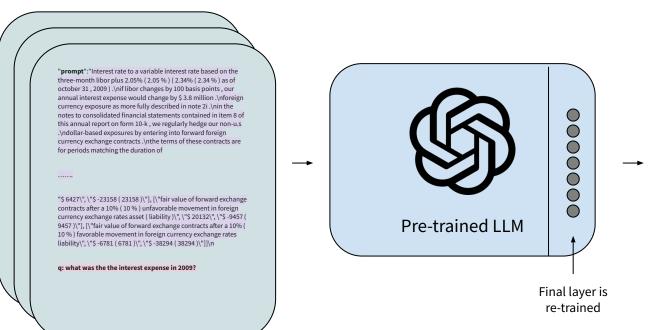
"\$ 6427\",\"\$ -23158 (23158)\"], [\"fair value of forward exchange contracts after a 10% (10%) unfavorable movement in foreign currency exchange rates asset (liability)\",\"\$ 20132\",\"\$ -9457 (\957)\"], [\"fair value of forward exchange contracts after a 10% (10%) favorable movement in foreign currency exchange rates

liability\", \"\$ -6781 (6781)\", \"\$ -38294 (38294)\"]]\n

g: what was the the interest expense in 2009?

Fine-tuning with pre-trained LLM

ModFinQA PROMPTS

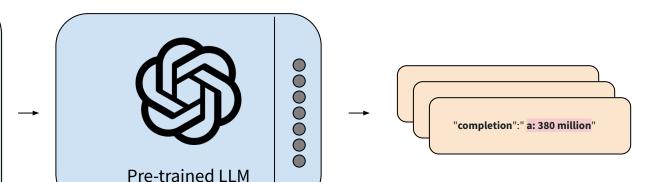


Fine-tuning with pre-trained LLM

ModFinQA PROMPTS

"\$ 6427\", \"\$ -23158 (23158)\"], [\"fair value of forward exchange contracts after a 10% (10%) unfavorable movement in foreign currency exchange rates asset [lability) \", \"\\$ 20132\", \"\\$ -9457 (9457) (9457) \", [\"fair value of forward exchange contracts after a 10% (10%) [avorable movement in foreign currency exchange rates liability\", \"\$ -6781 (6781)\", \"\$ -38294 (38294)\"]\")

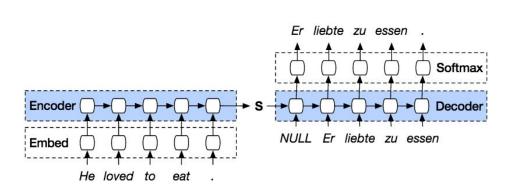
g: what was the the interest expense in 2009?

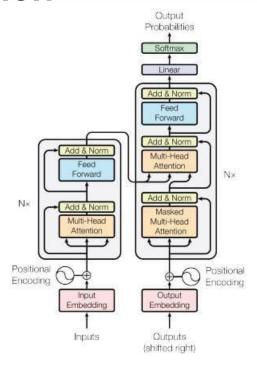


Final layer is re-trained

ModFinQA LABELS

Our Solution — Models/Evaluation





seq2seq

transformers

Our Solution — Models/Evaluation

- Initial benchmark model: seq2seq (**9%** accuracy)
- Improved model: fine-tuned t5-small (--% accuracy: generated empty tokens)
- Fine-tuned GPT-3 based model, more sophisticated, better performance (~55% accuracy)
- SOTA: FinQANet-Gold (RoBERTa-large): 68.76
- uman Performance: 50.68
- Peak (goated, expert, etc.) human performance: 91.16

Our Solution — Conclusion

- Fully neural-network based solution still suffers from hallucinations/miscalculations
- Probably best to combine these models with symbolic programs like Wolfram Alpha
- A good article: https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-w
 ork/
- But the remarkable—and unexpected—thing is that this process can produce text that's successfully "like" what's out there on the web, in books, etc. And not only is it coherent human language, it also "says things" that "follow its prompt" making use of content it's "read". It doesn't always say things that "globally make sense" (or correspond to correct computations)—because (without, for example, accessing the "computational superpowers" of Wolfram|Alpha) it's just saying things that "sound right" based on what things "sounded like" in its training material.