

Synthesizing an Instruction Tuning Dataset for Luxembourgish

Master Thesis

for the award of the academic degree
Master of Science (M.Sc.)

Universität Trier

FB II - Computerlinguistik und Digital Humanities
Natural Language Processing

Supervisor I: Prof. Dr. Achim Rettinger
Supervisor II: Dr. Keli Du

submitted on 04.11.2024 by:

Julian Valline
Kleeburger Weg 102
54296 Trier
s2juvall@uni-trier.de
1677068

Contents

Abstract	1
1 Introduction	2
2 Related Work	5
2.1 Creating Instruction Tuning Datasets	5
2.2 Multilingual Instruction Tuning Datasets	6
2.3 Synthesizing Instruction Tuning Data through Knowledge Distillation	6
3 Model comparison	8
3.1 Procedure	9
3.2 Empirical Tests	9
3.2.1 Questions/Instructions for comparing LLMs:	9
3.3 Assessment	15
3.3.1 Llama 3-8b-8192	15
3.3.2 Mistral-8x7b-32768	16
3.3.3 Gemma 2-9b-It	16
3.3.4 Llama-3.1-405b-instruct	17
4 Methodology	18
4.1 Data	18
4.1.1 Data structure	18
4.1.2 Filtering and Sampling	19
4.2 Instruction-Answer Pair Generation	20
4.3 Instruction Tuning Dataset	22
5 Experiments	24
5.1 Prompt Engineering	24
5.2 Synthetic Data Generation	30
5.2.1 Luxembourgish Instruction Tuning Dataset from Wikipedia	30
5.2.2 Luxembourgish Instruction Tuning Dataset from RTL news articles	32

5.3	Can Llama-3.1-405B-instruct answer it's own generated questions?	33
5.3.1	Wikipedia Data Questions	34
5.3.2	RTL News Questions	35
5.3.3	Model Question Answering Assessment	36
6	Evaluation	37
6.1	LLM evaluation	37
6.1.1	Sample data	38
6.1.2	Filtering Model Answers	38
6.2	Human evaluation	39
7	Conclusion	41
	Future Work	43
	Sworn Declaration	44
	Appendix	XIV
	Appendix A - Llama 3-8b-8192 model answers	XIV
	Appendix B - Mistral-8x7b-32768 model answers	XX
	Appendix C - Gemma2-9b-It model answers	XXXIV
	Appendix D - Llama 3.1-405b-instruct model answers	XL

Abstract

Pre-trained large language models (LLMs) have demonstrated noteworthy text generation capabilities. However, to execute convoluted, multi-turn user inquiries, instruction fine-tuning LLMs is pivotal. In this regard, we are the first to create a Luxembourgish instruction tuning dataset for fine-tuning state of the art LLMs. We use a large-scale, multilingual LLM to generate synthetic data from Luxembourgish language data. Our results show that our investigated multilingual 405B model is able to generate high-quality, synthetic instruction-answer pairs in Luxembourgish and can be applied at scale. We also show that smaller sized models have difficulties generating intelligible Luxembourgish language data. Finally, we provide a small-scale Luxembourgish instruction tuning dataset and a sampled version of it, which we evaluate with a 340B reward model and with human evaluators.

Chapter 1

Introduction

Recently, LLMs have shown to be exceptionally proficient at a variety of natural language tasks (Zhang et al., 2024; Brown et al., 2020; Touvron et al., 2023b; Yin et al., 2023; Peng et al., 2023; Li et al., 2023a). They have gained widespread and quick adoption as assistants (Shu et al., 2023), thanks to their facilitated usage through intuitive chat interfaces (Weber et al., 2024; Touvron et al., 2023b). These assistants are trained over several phases, starting with pre-training the model with a self-supervised learning objective, such as next-word prediction, on a sizable text dataset (Weber et al., 2024; Touvron et al., 2023b; Zhou et al., 2023). The incongruity between the training objective of next-word prediction and the users' aim to interact with an instruction-following, generative model is a significant problem with these pre-trained models (Zhang et al., 2024; Weber et al., 2024). Instruction tuning (IT) (Wei et al., 2022) overcomes this incongruity (Zhang et al., 2024) by teaching the model to interact with the user in a chat-like manner (Weber et al., 2024). Instruction tuning is the process of fine-tuning LLMs on large datasets, typically containing instruction-answer pairs. Given these instruction-answer pairs, the model is trained to follow the user's instruction by generating an appropriate answer to the provided instruction, which leads to more tractable and foreseeable model answers (Zhang et al., 2024). Additionally, the model's capacity to react to unknown instructions typically increases (Nayak et al., 2024; Zhou et al., 2023; Sanh et al., 2022).

Lately, open-access models, such as the Llama-3.1 series of models (Dubey et al., 2024), have caught up in performance to closed-sourced models (Labonne, 2024), like GPT-3 (Brown et al., 2020), raising the interest in researching open-sourced, instruction-tuned LLMs (Si et al., 2023). This allows for the cost efficient adaptation of open-access models to different applications through fine-tuning (Labonne, 2024).

Currently, there is a lack of open-source, non-English instruction tuning datasets for fine-tuning LLMs, as most IT datasets are primarily focused on English (Weber et al., 2024). This marginalizes other languages, as they lack in performance

in comparison and are also more costly to train and run (Weber et al., 2024). As far as we are aware, there is currently no instruction tuning dataset in Luxembourgish for fine-tuning state-of-the-art LLMs. In this work, we synthetically create an instruction tuning dataset in Luxembourgish from RTL news data and Luxembourgish Wikipedia articles. First, we compare 4 open-source, state of the art LLMs and test them for their Luxembourgish text generation capabilities (see chapter 3). We then select the best performing model, Llama-3.1-405B-instruct (Dubey et al., 2024), to synthetically generate 2420 instruction-answer pairs in Luxembourgish. We evaluate a randomly sampled subset of our instruction tuning dataset with Nemotron-4-340B-reward (Nvidia et al., 2024) and human evaluators. Our results show that Llama-3.1-405B-instruct is able to generate high-quality instruction-answer pairs in Luxembourgish, outperforming all other investigated LLMs, which struggle to generate intelligible Luxembourgish language data. Out of the 200 evaluated samples, 35% are of high-quality and can thus be used to build an instruction tuning set. Although we only generate a small sized dataset and evaluate a sub-sample of it due to limited resources, the model can generate an instruction tuning set at scale for fine-tuning large-sized, state of the art LLMs. We make the code and datasets publicly available¹.

Our key contributions are as follows:

- We create a novel, synthetically generated instruction tuning set in Luxembourgish, containing 2420 instruction-answer pairs in Luxembourgish.
- We provide a smaller-sized, evaluated instruction tuning set, subsampled from the aforementioned larger instruction tuning dataset in Luxembourgish.
- Our experiments on the small-sized dataset show that 35% of the data is of high-quality and can thus be used to build an instruction tuning set at scale.
- We compare 4 different open-source, state of the art LLMs for their Luxembourgish text generation capabilities and provide a detailed evaluation. Our experiments show that Llama-3.1-405b-instruct outperforms all other models and that smaller-sized models have difficulties generating comprehensible Luxembourgish language data.

¹<https://github.com/JulianValline/LuxInstructTuning>

This thesis is structured as follows: Chapter 2 provides a literature review of related work, encompassing the creation of instruction tuning datasets, multi-lingual instruction tuning datasets, and the synthesis of instruction tuning data through knowledge distillation. Chapter 3 undertakes a comparative analysis and evaluation of four distinct open-source language models, focusing on their capabilities in comprehending and generating Luxembourgish text. Chapter 4 details the methodology employed for the creation of the instruction tuning dataset. Chapter 5 describes the experiments conducted, including prompt engineering for data generation from Wikipedia and RTL news articles, and an evaluation of the model's ability to answer its own generated questions. Chapter 6 presents an evaluation of the resulting dataset, and chapter 7 concludes the thesis with a summary of findings and potential future research directions.

Chapter 2

Related Work

2.1 Creating Instruction Tuning Datasets

An instruction tuning dataset is typically composed of instructions, inputs and outputs (Zhang et al., 2024). Instructions are natural language descriptions of the task the model is supposed to fulfill (Zhang et al., 2024). Inputs are the data that the model should process in accordance with the instructions (Zhang et al., 2024), ranging from simple questions to more complex instructions. In light of the instruction and input, outputs are the desired answers or completions that the model is expected to produce (Zhang et al., 2024).

According to Zhang et al. (2024), instruction tuning datasets are build primarily in two different ways:

- **Data integration** combines and structures data from various annotated datasets into a suitable, unified format, such as instruction-answer pairs (Zhang et al., 2024), resulting in datasets like Flan (Longpre et al., 2023) and P3 (Bach et al., 2022).
- **Data generation** uses LLMs to synthetically generate outputs based on the provided instruction (Zhang et al., 2024), resulting in datasets like Self-Instruct (Wang et al., 2023), where instructions, inputs and outputs are all obtained through a LLM.

Our work leans on the data generation approach; we use a LLM to generate the instructions and expected answers, based on a handcrafted prompt fed to the model to generate instruction-answer pairs from a short text snippet (See chapter 4).

2.2 Multilingual Instruction Tuning Datasets

Extending on P3, xP3 (Muennighoff et al., 2023) is a multilingual, human-crafted instruction tuning dataset, where the data is taken from P3 and other multilingual datasets (Zhang et al., 2024). The dataset is constructed following the aforementioned data integration method, where data from various sources is combined into a unified format (Zhang et al., 2024). Muennighoff et al. (2023) further extend xP3 to xP3mt by applying machine translation (Weber et al., 2024).

Multilingual datasets like Bactrian-X (Li et al., 2023b) emerged from machine-translated Alpaca instructions and matching, GPT-3.5-Turbo (Brown et al., 2020) generated answers (Weber et al., 2024). LIMA (Zhou et al., 2023) is a fine-tuned version of Llama (Touvron et al., 2023a), trained on 1000 attentively selected instruction-answer pairs (Zhang et al., 2024; Weber et al., 2024). Weber et al. (2024) extend LIMA into Lima-X by translating the instructions from the original LIMA dataset into 4 languages.

2.3 Synthesizing Instruction Tuning Data through Knowledge Distillation

LLMs are used to generate synthetic instruction tuning datasets, benefiting over handcrafted IT sets in both, quality and resource efficiency (Zhang et al., 2024). In knowledge distillation, a more compact model, called the "student" model, is trained to imitate the actions of a more sophisticated, larger model, called the "teacher" model (Zhang et al., 2024). Knowledge distillation is primarily used to transfer knowledge from the teacher model to the student model, enabling the student model to perform comparably despite being much smaller and more effective (Zhang et al., 2024).

Zhang et al. (2024) thus describes the creation of synthetic instruction tuning sets through knowledge distillation by generating instructions using a large "teacher" model like GPT-3 (Brown et al., 2020) and fine-tune a smaller "student" model like Llama-7B (Touvron et al., 2023a) to follow the larger model's instructions, as illustrated by Taori et al. (2023). They developed Alpaca (Taori et al., 2023), a modified Llama-7B model fine-tuned on 52k instruction-input-response triplets (Zhang et al., 2024), which are generated with text-davinci-003 (Brown et al., 2020). Alpaca performed on par with text-davinci-003 from OpenAI (Zhang et al., 2024).

Peng et al. (2023) use GPT-4 (OpenAI et al., 2024) to generate the output for the

instruction-input pairs used in the original Alpaca dataset. Our work follows a similar dataset structure as Peng et al. (2023), except we use Llama-3.1-405B-instruct to generate all new instructions and expected answers in Luxembourgish.

Chapter 3

Model comparison

To best evaluate LLMs on their Luxembourgish text generation capabilities, we need to create a high-quality instruction tuning dataset for Luxembourgish. Doing this manually, requires too many resources and cannot compete quality wise with LLM generated output. Li et al. (2023b) use GPT-3.5-turbo to generate instruction tuning sets in different languages.

As previously mentioned, open-access models have narrowed the performance gap on closed-access models (Labonne, 2024; Si et al., 2023). Open-access models, such as Llama, have the advantage that their produced data can be freely used to fine-tune other LLMs and redistribute them¹. Hence, our focus lies on open-access models.

In this chapter, we compare 4 open-access, state of the art LLMs for their Luxembourgish capabilities and evaluate them based on accuracy, grammar, fluency and comprehensiveness. We choose Llama3-8b-8192 (Dubey et al., 2024), Gemma2-9b-It (Team et al., 2024b), Mistral-8x-7b-32768 (Jiang et al., 2024) and Llama-3.1-405b-instruct. We find that Llama-3.1-405b-instruct outperforms all other models in understanding and generating Luxembourgish text. We also briefly tested Nvidia's Nemotron-4-340b-instruct (Nvidia et al., 2024), but we didn't find any noticeable differences compared to Llama3, Gemma2 and Mistral-8x, so we omit it from this section. All 8b and 9b models were tested on Groq² while Llama-3.1-405b was tested directly on Nvidia NIM³. We describe our experiments in the following.

¹Llama-3.1-Community License available at: https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/LICENSE

²<https://groq.com/>

³https://build.nvidia.com/explore/discover#llama-3_1-405b-instruct

3.1 Procedure

We test each model with a set of tasks across different domains, such as language understanding, sentiment analysis, translation, text generation, error detection, conversational dialogue, question answering (in and about Luxembourg), summarization and extractive question answering. For the last part, we sample from Luxembourgish Wikipedia and RTL news articles from May 2024. We prompt the model to respond to each task with a clear and concise answer. For tasks that require translation or text generation, the model is instructed to provide a written response. For tasks that require sentiment analysis or error detection, we ask the model to provide a sentiment label or a suggested correction. The exact instructions can be found in section 3.2.

3.2 Empirical Tests

The following section contains the exact instructions that are given to the model. We try both English and Luxembourgish instructions for some tasks. The different model answers can be found in Appendix A, B, C and D respectively.

3.2.1 Questions/Instructions for comparing LLMs:

Initial instruction:

Please respond to each task with a clear and concise answer. For tasks that require translation or text generation, please provide a written response. For tasks that require sentiment analysis or error detection, please provide a sentiment label or suggested correction.

Task 1: Language Understanding

Paraphrase the following sentence in Luxembourgish: Lëtzebuerg ass eng kleng Staat.

We evaluate the models' ability to understand the original sentence and generate a paraphrased version.

Task 2: Sentiment Analysis

1. Analyze the sentiment of the text and respond with a sentiment label (e.g. positive, negative, neutral): D'Wirtschaft vum Land ass momentan schlecht.
2. Analyze the sentiment of the text and respond with a sentiment label (e.g. positive, negative, neutral): Ech hunn dech gär, mee mir passen net zesummen.
3. Analyze the sentiment of the text and respond with a sentiment label (e.g. positive, negative, neutral): Ech hunn haut 300€ am Lotto gewonnen.
4. Analyze the sentiment of the text and respond with a sentiment label (e.g. positive, negative, neutral): Ech hunn haut 300€ op der Strooss verluer.
5. Analyze the sentiment of the text and respond with a sentiment label (e.g. positive, negative, neutral): Ech hunn haut mäin Brudder getraff.

We evaluate the models' ability to accurately identify the sentiment of the text.

Task 3: Translation

1. Translate the following sentence into Luxembourgish: The weather is nice today
2. Iwwersetzt dëse Saz op Lëtzebuergesch: "The European Union has its headquarters in Brussels."
3. Iwwersetzt folgend Saz op Englesch: "Ech ginn an de Park."
4. Iwwersetzt dëse Saz op Lëtzebuergesch: I'm going to the park
5. Iwwersetzt dësen Text op Lëtzebuergesch: "Luxembourg is a small country in Europe."

Task 4: Text Generation

1. Beschreif d'Lëtzebuergesch Kultur. generate a short text (e.g. 1-2 sentences) in response

2. Write a short poem in luxembourgish about the beauty of the Luxembourgish Moselle region
3. List five Luxembourgish idioms
4. Explain the Luxembourgish tradition of "Bretzelsonndeg" in Luxembourgish
5. Generate a paragraph of text in Luxembourgish about the history and culture of Luxembourg.

We evaluate the model's ability to generate text in a grammatically and factually correct way

Task 5: Error Detection

Den Letzebuergeschen Land huet 2000 Awunner an ass méi grouss als Daitschland. identify the errors and suggest a correction (grammar and content).

We evaluate the models' ability to detect and correct errors

Task 6: Conversational Dialogue

Wéi géifs du engem Tourist Lëtzebuerg an dräi Sätz beschreiwen.

We evaluate the models' ability to respond coherently and relevantly to your questions and statements.

Task 7: Luxembourgish Questions

1. Wien ass de Premierminister vu Lëtzebuerg?
2. Wéi eng Sproochen gi geschwat zu Lëtzebuerg?
3. Wat ass d'Hauptstad vu Lëtzebuerg?
4. Wéi vill Awunner huet Lëtzebuerg?
5. Wéi ass d'Wieder haut zu Lëtzebuerg?
6. Wie war de Groussherzog vu Lëtzebuerg am Joer 2000?

Task 8: Summarization

Summarize the following Luxembourgish text into a few short sentences in Luxembourgish:

Lëtzebuerg ass eng representativ Demokratie an eng konstitutionell Monarchie mat engem Groussherzog als Staatschef, wouduerch d'Land dat eenzegt Groussherzogtum ass, dat nach ëmmer als souveräne Staat op der Welt existéiert. Seng dynamesch Wirtschaft mécht et zu engem vun de räichsten a räichste Länner vun der Welt, mat engem vun den héchste Bruttoinlandsprodukter pro Kapp op der Welt. D'Wirtschaft konzentréiert sech haaptsächlech op finanziell Aktivitéiten (ongeféier d'Halschent vum Bruttoinlandsprodukt), favoriséiert duerch ganz avantagéis Besteuerung a bestëmmte Beräicher (virtuell Steuerbefreiung fir Gewinn aus der Exploitatioun vu Patenter oder Software). Der ONG Oxfam no (2017 Zuelen) gehéiert Lëtzebuerg zu "attraktiveste Steuerparadäiser, déi vun de Firme benotzt ginn", eng Situatioun déi sech zënterhier e bësse geännert huet. Déi zentral Lag vum Lëtzebuerger Territoire an Europa huet et historesch eng Plaz vu grousser strategescher Wichtigkeet fir vill Muechten gemaach, zanter sengem Ufank als Hinterland vum Réimesche Rheinland, an de Bau, op de Ruine vun engem fréiere réimesche Fort, eng feudale Buerg. , am postkarolingesche Kontext am Héichmëttelalter, duerno seng strategesch Lag am Südoste vun de Groussnidderlanden, tëscht dem 15. an dem 17. Joerhonnert successiv burgundesche a spuenesche.

Text taken from Wikipedia article about Luxembourg in Luxembourgish⁴

Task 9: : Extractive Question Answering:

1. Generate 5 questions and answers in Luxembourgish from the following text:

Um Wiener Kongress 1815 gouf dat viregt Herzogtum Lëtzebuerg vun den europäesche Groussmuechten net "restauréiert". D'Preisen, déi d'Ostufervun der Musel scho besat haten, kruten nach all déi al lëtzebuergesch Gebidder ëstlech vun der Our, mat Ausnam vu Veianen, an déi ëstlech vun der Sauer. Aus dem Gebitt vum Département des Forêts an Deeler vum Département Sambre-et-Meuse ass, aus reng strateegesch-diplomateschen

⁴[https://lb.wikipedia.org/wiki/L%C3%ABtzebuerg_\(Land\)](https://lb.wikipedia.org/wiki/L%C3%ABtzebuerg_(Land))

an dynastesche Grënn, ee Groussherzogtum ginn. De Wëllem vun Oranien-Nassau, dee grad kuerz virdrun als hollännesche Kinnek Wëllem I. unerkannt gi war, gouf a Personalunioun Groussherzog vu Lëtzebuerg. Dat neit Lëtzebuerger Land gouf och an den Däitsche Bond (eng Zort Ersatzräich) integréiert, an d'Festung Lëtzebuerg gouf eng "Bundesfestung", wat dem preisesche Kinnekräich d'Garnisounsrecht erméiglecht huet. Vun Zäit zu Zäit sollten d'Vertrieder vun deene ronn 40 Länner, déi am Däitsche Bond waren, sech ënner dem Virsätz vum Keiser vun Éisträich zu Frankfurt treffen. Well de Wëllem I., dee Kinnek vun de Vereenegten Nidderlande war, och als Herrscher vu Lëtzebuerg wéi ee Kinnek wollt behandelt ginn, huet hien um Wiener Kongress drop gehale fir "Grouss"-Herzog ze ginn, also eng "Kinneklech" Houheet (deut. "Königliche Hoheit", fra. "Altesse royale"), wat en als "einfachen" Herzog vu Lëtzebuerg net hätt kënnen sinn. Dat ass d'Erklärung, firwat Lëtzebuerg als GroussHerzogtum nei entstanen ass (et huet deemno glat näischt mat der Gréisst vum Land ze dinn). 1830 huet dem Wëllem I. seng onglécklech-ongerecht Politik a seng Södprovënzen d'Belsch Revolutioun ausgeléist, déi och ee gudden Deel vun de Lëtzebuerger begeeschtert huet, well hien, eise Groussherzog, säi Groussherzogtum praktesch ëmmer nëmme wéi déi 19. Provënz vu sengem Räich behandelt hat. Den 19. Abrëll 1839, néng Joer no der Grënnung vun der Belsch, gouf um Traité vu London, no laangem hin an hier, d'Groussherzogtum an zwee gedeelt: de westlechen Deel, deen haaptsächlech franco-romanesch (Wallounesch a Louthrengesch) war, gouf als Provënz Lëtzebuerg der Belsch zougesprach, an de germanophonen, also östlechen Deel, ass - weiderhin als Groussherzogtum - fir d'éischt Kéier een eegestännege Staat ginn, allerdéngs nach ëmmer (bis 1890) a Persunalunioun mat Holland.

Text taken from Wikipedia article about Luxembourg in Luxembourgish⁵

2. Generate 3 questions and answers in Luxembourgish from the following article:

D'Wourecht iwwer säi bescht Stéck!

Wéivill Fraen-Zäitschrëften, wéivill Internet-Portaler, wéivill Jugend-Magaziner hu sech scho mat der Gréisst vum Mann sengem beschte Stéck beschäftigt? A wahrscheinlech huet och all Mann sech schonn emol d'Fro gestallt, wéi wichteg d'Gréisst beim Geschlechtsverkéier ass. Vill gëtt gezielt, ma d'Realitéit ass ganz einfach: Dat wat zielt ass d'Zesummespill vu béide Partner. Op wat kënnt et dann elo wierklech un? E groussen Deel vum

⁵[https://lb.wikipedia.org/wiki/L%C3%ABtzebuerg_\(Land\)](https://lb.wikipedia.org/wiki/L%C3%ABtzebuerg_(Land))

Sex spillt sech am Kapp of. Wee sech mat Leidenschaft op deen anere freet, ween et kaum nach kann erwaarden, dee ka bal net enttäuscht ginn. Wichteg ass, datt ee mat vill Gefill op säi Partner ageet. D'Ziel ass ze verstoen, wat deem anere Pleséier mécht. Natierlech spillen och d'Technik an d'Positioun eng grouss Roll, well net all Stellung de selweschten Effekt huet. De Mythos iwwert d'Gréisst bréngt awer virun allem eppes mat sech: datt vill Männer doduerch immens ënner Drock geroden. A genee duerch deen Drock kann et hinne schwéier falen hir Sexualitéit richteg auszeliewen an ze genéissen. Wahrscheinlech ass dëst och de Grond firwat ëmmer méi Männer sech ënner d'Messer leeën. Mä och eng Verlängerung ännert net onbedéngt eppes. Schliisslech ännert déi näischt un der Sensibilitéit géigeniwwer der Partnerin. Dat eenzegt wat hëlleft ass, wann een op seng Partnerin an op där hir Kierpersprooch lauschtert.

Text taken from the RTL News articles, May 2024 (Sample 0)

3. Generate 3 questions and answers in Luxembourgish from the following article:

Mariah Carey – Glitter

Zanter 8 Deeg ass dëm Mariah Carey säin neisten CD an dë Butteker zë fannen. 'Glitter' heescht den Album, deen d'selwécht heescht ewéi dee Film an deem d'Mariah Carey säin Début am Filmbusiness mécht. Mat deem CD gët et och eng nei Single. An do kann een opootmen, d'Mariah Carey quitscht nët méi op daer neier Single, mee et gët sech all Méi fier déi verschidde Facetten vu senger Stëmm zë weisen, déi ëmmerhin iwwert 8 Octaven geet. 'Never Too Far' heescht déi neist Auskopplung aus sengem aktuellen CD, an domat geet déi amerikanesch Musekerin nees zëreck un d'Originen vun hierer Carrière, mat enger Ballade déi sech an d'Traditioun vun hieer gréisste Succès'en 'à la' Vision Of Love', 'Emotions', 'One Sweet Day' a 'Without You' areit. 'Loverboy' an 'Don't Stop', déi zwou éischt Single'en aus dësem neien CD kann een deemno rouëg vergiessen, well ausser gequitschs an engem Videoclip an deem d'Mariah Carey an engem hautenken Top op dë Stroossen danzt haaten déi zwee Lidder nët vill zë bidden, och wann se vun hierem Opbau hier schon dë 'Retour aux Sources' ugëkënnegt hun. Dë Sound vun dën 80er ass op 'Glitter' omniprésent, dën Album soll eng musekalësch Rees duërch dëm Billie säi Geescht duerstellen. D'Billie ass dë Personnage deen d'Mariah Carey am Film 'Glitter' spillt. Et ass also um bäschte Wée, fier säin Abonnement op d'Spëtze positiounen an dën Hitparaden zë verlängeren. Well d'Mariah Carey ass deen eenzegen Artist deen et färdëg brüecht

huët während dën 90er Joren all Joër eng Nummer 1 an dën Charts zë placéieren. Doniewt haat hatt ësouvill Nummer 1 Hits wéi keen aaneren weiblëchen Artist fierdrun, an nëmmen d'Beatles an dën Elvis Presley kënnen ët nach iwwerbidden. A mat Auszeechnungen bei dë Grammy'en, dën American Music Awards, dë Billboard Music Awards an dë World Music Awards kann ee sëch weise loossen. Niewt dër Musek wëll d'Mariah Carey ëlo och am Filmbusiness Fouss faassen. 'Glitter' heescht deen éischte Film deen hatt gëdréit huët, an do spillt d'Sängerin d'Billie, ee jonkt Meedchen, daat nët driwwer ëwechkennt, dass ët als Bebé ausgesaat gin ass. Doniewt huët ët och schon d'Dréiaarbëchten fier dën independent Film 'Wise Girls' mam Mira Sorvino oofgeschloss.

Text taken from the RTL News articles, May 2024 (Sample 1800)

3.3 Assessment

All model responses are evaluated by a Luxembourgish native speaker on the following criteria:

- **Accuracy:** How accurate are the model's responses to factual questions about Luxembourgish language, culture, history, and geography?
- **Grammar:** Is the grammar of the generated text correct?
- **Fluency:** How natural and fluent does the model's Luxembourgish language sound?
- **Comprehensiveness:** Does the model provide sufficiently detailed and informative answers?

3.3.1 Llama 3-8b-8192

The model has difficulties understanding the meaning of Luxembourgish sentences in some tasks. In the language understanding task for example, it does not recognize the meaning of the word "staat" as being the country of Luxembourg but rather thinks of it as the city of Luxembourg. For translation tasks, it seems to be better at translating from Luxembourgish to English than the other way around. When translating into Luxembourgish, it often gets a wrong translation or just wrong articles. The text generation task shows the model's inability

to generate semantically or factually correct Luxembourgish text. The poem makes absolutely no sense and uses wrong articles, it only gets 2 out of 5 idioms correctly and gets only partly correct content for the paragraph generation about the history of Luxembourg. The error detection seems to work to a certain extent. It can correct simple content but no grammar. The simple question answering task seems to work best, as the content and grammar are correct, and it only makes little grammatical mistakes. However, we assume that this works best given that these are very common questions about Luxembourg so the model might have learned to answer these specific questions. Overall, the model seems to have a lot of difficulties when it comes to grammar and the use of correct Luxembourgish words. Across all tasks, the model either makes up some words or takes them from other languages, such as German or French. We provide the model answers in Appendix A.

3.3.2 Mistral-8x7b-32768

The model does not seem to understand the meaning of Luxembourgish sentences in some tasks. Compared to Llama 3, the model cannot paraphrase the sentence in the language understanding task. Sentiment Analysis seems to work well again. In addition to Llama 3, Mistral also adds an explanation for each sentiment. For the translation task, the model again only translates correctly from Luxembourgish to English but not the other way around. The model makes up some words but less than Llama 3. The prepositions are also often wrong. Mistral also has difficulties with text generation. The Luxembourgish text is grammatically wrong, content is partly wrong, the poem is totally wrong in terms of grammar and semantics and the idioms are made up. Analogous to Llama 3, the model struggles with grammar overall and the use of correct Luxembourgish words. Furthermore, many Luxembourgish questions are simply answered in English rather than in Luxembourgish. We provide the model answers in Appendix B.

3.3.3 Gemma 2-9b-It

The model seems to be slightly better than Llama 3 and Mistral grammatically speaking. In the language understanding task, the model is able to paraphrase the sentence even if it is a little farfetched. Sentiment analysis also works well, and it provides additional explanations to why sentiments are labeled that way. However, it deduces a wrong context with one word. The translation does not

work well, sentences get wrongly translated and the grammar is off. Some words are completely made up and do not exist in Luxembourgish. The text generation is rather general and simple in meaning and not grammatically correct, some words are completely wrong and are mostly taken from German. The poem is semantically wrong, words are taken from English or just made up. The idioms also did not work at all. The error detection only works on content but not on grammar. Simple question answering only works to some extent. Some questions are answered incorrectly. Overall, the model works slightly better than Llama 3 and Mistral 8 but it still makes lots of grammatical mistakes and uses made up words. We provide the model answers in Appendix C.

3.3.4 Llama-3.1-405b-instruct

The model correctly paraphrases the sentence in the language understanding task. It even adds more context. The grammar is also correct. Sentiment analysis also works great, adding additional explanation to the labels. The translation task also works well, correctly translating all sentences from Luxembourgish to English and the other way around. The text generation task is mostly correct. The generated answer is broad and general, but the grammar and content are correct, except one fact, where it says that Luxembourg is bilingual, however there are three official languages. The poem also does not work well with this model. The content makes little sense, and it still makes up some words. The idioms are also made up, they are grammatically incorrect and there are very few made up words. When asking the model to explain the Luxembourgish tradition of “Bretzelsonndeg”, the model returns a redundant answer being mostly grammatically correct. There were again very few made up words. The paragraph about the history and culture of Luxembourg on the other hand was really good, grammatically and content wise. Error detection also works well. It corrects the sentence’s content and grammar. The conversational dialogue and question answering tasks were also correct. Furthermore, the model does well on text summarization and extractive question answering, which is really important, since the latter is going to be our actual use case for creating the instruction tuning dataset. Given two articles, taken from the RTL new articles from May 2024, the model is able to generate question and answer pairs, only making minor mistakes, such as the wrong grammatical tense in 2 sentences. Overall, Llama 3.1-405b-instructs outperforms all other models by far in terms of accuracy, grammar, fluency and comprehensiveness. We provide the model answers in Appendix D.

Chapter 4

Methodology

The following chapter describes our procedure for creating the instruction tuning dataset.

4.1 Data

We have two main data sources, the entirety of the Luxembourgish Wikipedia and all RTL news articles and comments up to May, 2024. We download the latest version of the Luxembourgish Wikipedia dump¹. We then extract and clean the Wikipedia dump using Wikiextractor². Finally, we convert the extracted data into a JSON file for facilitated handling. The RTL news articles are cleaned and also transformed into JSON format.

This results in 1.2 million RTL comments, 303k RTL news articles and about 80k Wikipedia articles, totalling up to 1.6 million data entries. We decide to omit RTL comments, as they are short and may contain biased opinions.

4.1.1 Data structure

Both data sources are contained in a JSON file. Figure 4.1 shows an example entry of the RTL news articles. We are mainly interested in the `public_date`, `title`, `header` and `text`, as these contain the most relevant information about the news article. The Wikipedia data has a similar structure. Here we are interested in `title` and `text` only. We show an example entry in Figure 4.2.

¹<https://dumps.wikimedia.org/lbwiki/latest/>

²<https://github.com/attardi/wikiextractor/>

```
{'category_name': 'International',
 'category_id': 5,
 'article_id': 2190645,
 'type': 'news',
 'public_date': '2024-04-28 14:04:54',
 'title': '"USA sinn eenzegt Land, dat Israel dorun
hënnere kann, Rafah unzegräifen"',
 'header': 'Dat sot de Palästinenserpresident Mahmud Abbas
op enger Reunioun um Weltwirtschaftsforum a Saudi-Arabien.
',
 'text': ' Den Abbas huet d\'US-Regierung ganz däitlech
dozou opgeruff, alles ze maachen, fir dass eng israeelesch
Buedemoffensiv op Rafah am Süde vun der Gazasträif
verhënnert gëtt... ',
 'tags': ['international'],
 'text_id': 301779,
 'lang_id': 'lb'}
```

Figure 4.1: Example RTL News entry in JSON format. Text shortened for display reasons.

```
{'id': '25',
 'revid': '580',
 'url': 'https://lb.wikipedia.org/wiki?curid=25',
 'title': 'Matthew Perry',
 'text': 'De Matthew Langford Perry, gebuer den 19.
August 1969 zu Williamstown am Massachusetts, a
gestuerwen den 28. Oktober 2023 zu Los Angeles, war
en US-amerikanesch-kanadesche Schauspiller, dee virun
allem duerch seng Roll als "Chandler Muriel Bing" an der
Televisiounsserie "Friends" bekannt ginn ass...'}

```

Figure 4.2: Example Wikipedia article in JSON format. Text shortened for display reasons.

4.1.2 Filtering and Sampling

Given our limited available resources, we cannot work with the full amount of data at our disposal. The synthetic data generation would simply be too costly. Our idea is to prompt the model to create 5 instruction-answer pairs for each

news article or Wikipedia article, which would increase the data amount five-fold. Thus, we decide to randomly sample 100 Wikipedia articles and 400 RTL news articles from our data pool, totalling up to 500 examples, which would result in a dataset of 2500 entries after generating the data. Our intuition is, that Llama-3.1-405b-instruct has likely already seen most Luxembourgish Wikipedia articles, hence we decide to include only 100 examples into our dataset and sample the rest from the RTL news articles, which the model probably has not yet seen. We still want to include Luxembourgish Wikipedia articles, as they contain general world knowledge in Luxembourgish, which is crucial for instruction tuning. RTL news articles are more Luxembourg oriented, containing valuable knowledge about Luxembourg. Additionally, we only sample from articles that contain at least 1500 words, ensuring that the articles contain enough information to facilitate the generation of 5 unique question-answer pairs for each article.

4.2 Instruction-Answer Pair Generation

We use Llama-3.1-405b-instruct to create synthetic instruction-answer pairs in Luxembourgish given the provided RTL news or Wikipedia article. As aforementioned, we only use 100 samples from the preprocessed Wikipedia dump and 400 samples from the RTL news data, to keep the dataset small. We create 5 instruction-answer pairs for each entry, totalling up to 2500 synthetic instruction-answer pairs in Luxembourgish.

It is crucial, to feed the best possible instruction prompt to the language model to ensure high quality data generation in the desired output format. We try to define a very detailed and accurate instruction, covering every possible case within our data. We also try to include the news article date in the instruction and answer formulation. Finally, we want the model to return the data as a JSON file, putting every 5 question-answer pairs for each article in a separate file. The Prompt Engineering experiments for this step are shown in detail in Section 5.1. The instruction is formulated as follows:

"The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning. You are an assistant and generate all your responses in Luxembourgish. You will receive a series of texts one after the other. Generate a dataset from the text given with 5 unique pairs of instruction and expected response. Give elaborate and high-quality responses and answer as accurately as possible. Make sure that the instruction contains all relevant information that

is needed to give the response. If the question does not contain enough information to be answered accurately, reply that you need more information to answer the question. Whenever a date is provided for a news article, include the date in the instruction and its respective response or just in the response, depending on what you consider best in the context. Also give temporal context if possible. Name the columns instruction and response respectively. Put everything into a JSON file."

The data is fed to the model in the following way: For each RTL news article, we combine the instruction, date, title, header and the news article text.

For the Wikipedia articles, we follow a similar structure, except we only combine the instruction, the title and the text. We show the structure of the model prompts in Figure 4.3 and Figure 4.4 respectively.

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning...

Text:

News article from: 2018-12-07

Mat Summerpneuen en Accident gehat an ugedréckt

Bei -4 Grad war e Chauffeur am Februar zu Bieles
accidentéiert. Hien hat Summerpneuen drop a war duerno
fortgaangen ... an hofft elo op eng niddreg Strof.

D'Zäit vun de manner gudde Fuerkonditiounen op eise
Stroosse réckelt am Prinzip méi no. An deem Fall, ...

Figure 4.3: Example Prompt for an RTL News Article fed to the LLM. We truncate the text for display reasons.

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning...

Text:

Leo VIII. (Poopst)

De Leo VIII., gestuerwen den 1. Mäerz 965 zu Roum, war vum 4. Dezember 963 bis zu sengem Doud, Poopst vun der Kathoulescher Kierch. Säi Poopstnumm bedeit: "de Léiw" (Latäin.)...

Figure 4.4: Example Prompt for a Luxembourgish Wikipedia article fed to the LLM. We truncated the text for display reasons.

4.3 Instruction Tuning Dataset

We follow a similar structure as the dataset released by Peng et al. (2023)³. Akin to Peng et al. (2023), our dataset is composed of three columns: instruction, response and text. The columns instruction and response contain the generated instructions and matching answers respectively. The column text contains a prompt template, combining the two columns. We show the structure of our prompt template in Figure 4.5 and the structure of our instruction tuning dataset in Table 4.1.

The following provides an instruction in form of a question. Give a response that answers the question accurately.

###Instruction: {Instruction}

###Response: {Response}

Figure 4.5: Prompt Template combining instruction and response columns of the instruction tuning dataset.

³<https://huggingface.co/datasets/vicgalle/alpaca-gpt4?row=0>

instruc- tion	response	text
Ginn déi wichtegst Informatiounen iwwer d'Deborah De Robertis.	D'Deborah De Robertis ass eng lëtzebuergesch-franséisch Performance-Kënschtlerin, gebuer den 12. Februar 1984 zu Lëtzebuerg. Si ass um Lampertsbiërg opgewuess an huet op der École de recherche graphique zu Bréissel studéiert.	<p>The following provides an instruction in form of a question. Give a response that answers the question accurately.</p> <p>###Instruction: Ginn déi wichtegst Informatiounen iwwer d'Deborah De Robertis.</p> <p>###Response: D'Deborah De Robertis ass eng lëtzebuergesch-franséisch Performance-Kënschtlerin, gebuer den 12. Februar 1984 zu Lëtzebuerg. Si ass um Lampertsbiërg opgewuess an huet op der École de recherche graphique zu Bréissel studéiert.</p>
Wat ass déi bekanntst Performance vun der Deborah De Robertis?	Déi bekanntst Performance vun der Deborah De Robertis ass 'Miroir de l'origine', wou si sech viru Konschtwierker a Muséeën ouni Bëxelche fotograféiere gelooss huet, als Referenz op dem Gustave Courbet säin Tableau 'L'origine du monde'.	<p>The following provides an instruction in form of a question. Give a response that answers the question accurately.</p> <p>###Instruction: Wat ass déi bekanntst Performance vun der Deborah De Robertis?</p> <p>###Response: Déi bekanntst Performance vun der Deborah De Robertis ass 'Miroir de l'origine', wou si sech viru Konschtwierker a Muséeën ouni Bëxelche fotograféiere gelooss huet, als Referenz op dem Gustave Courbet säin Tableau 'L'origine du monde'.</p>

Table 4.1: Instruction Tuning Dataset Structure. We show the first 2 rows of the dataset.

Chapter 5

Experiments

In this chapter, we describe our experiments done for generating the Luxembourgish instruction tuning set.

5.1 Prompt Engineering

An important step for attaining high quality data is to prompt the LLM with the best possible prompt by carefully crafting and refining instructions fed to the model. This procedure is called prompt engineering. In the following, we will go through our most important prompt engineering steps and intermediate results. All prompt engineering experiments were performed on Llama-3.1-405b-instruct through Nvidia NIM¹. We show our first version in Figure 5.1 and the respective model output in Figure 5.2:

¹<https://www.nvidia.com/en-us/ai/>

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning. You are an assistant and generate all your responses in Luxembourgish. You will receive a series of texts one after the other. Generate a dataset from the text given with 5 unique pairs of instruction and expected response. Name the columns instruction and response respectively. Give high quality responses. Put everything into a JSON file.

Text:

<Title>

<Article>

Figure 5.1: 1st version of the instruction prompt

```
[
  {
    "instruction": "Wie viele Kandidaten haben sich
f\u00fcr die Komplementarwahlen in der Fusiounsgemeng
Bous-Waldbriedemes gemeldet?",
    "response": "7 Kandidaten hu sech gemellt."
  },
  {
    "instruction": "Wer ist der Buergermeeschter vun
der Ostgemeng?",
    "response": "Den Toto Da Costa Araujo ass zanter
zejoert Buergermeeschter."
  },
  ...
]
```

Figure 5.2: Generated instruction-answer pairs in JSON format given our 1st version of the instruction prompt. We shorten the output due to display reasons.

This is a solid baseline but we noticed that the generated instruction-answer pairs were not of high quality. We need to ensure that all generated questions have a unique answer. In some examples, the answer contained information which was not clearly or specifically requested by the instruction. Hence, we added to the prompt that the question needs to include all needed information

to formulate a unique answer. As an anti-hallucination measure, we also added to the prompt that the answer should specifically mention that it needs more information if said information is not sufficiently provided by the instruction. This resulted in the 2nd version of our instruction prompt, as shown in Figure 5.3 with respective output in Figure 5.4.

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning. You are an assistant and generate all your responses in Luxembourgish. You will receive a series of texts one after the other. Generate a dataset from the text given with 5 unique pairs of instruction and expected response. Give high quality responses and answer as accurately as possible. Make sure that the response contains all relevant information that is needed to give a clear answer. If the question does not contain enough information to be answered accurately, reply that you need more information to answer the question. Name the columns instruction and response respectively. Put everything into a JSON file.

Figure 5.3: 2nd version of the instruction prompt

```
[
  {
    "instruction": "Wie ass de Grond fir
d'Komplementarwalen an der Fusiounsgemeng
Bous-Waldbriedemes?",
    "response": "De Grond ass d'Demissioun vum
Conseiller Mike Molling wéinst Incompatibilitéits-Grënn
mat sengem
Fonctionnaires-Beruff."
  },
  {
    "instruction": "Wéi vill Kandidaten hu sech fir
d'Komplementarwalen an der Sektoun Waldbriedemes
gemellt?",
    "response": "Et hu sech 7 Kandidaten gemellt."
  },
  ...
]
```

Figure 5.4: Generated instruction-answer pairs in JSON format given our 2nd version of the instruction prompt. Output shortened due to display reasons.

This already looked better but then we realized the importance of the news article date for RTL news when generating the data. Hence, we included the news article date explicitly into the instruction prompt, resulting in Figure 5.5 and Figure 5.6. We noticed that some news article dates were invalid, like having year 0000 and others had future dates like 2026. The LLM was able to automatically filter out the invalid dates when generating the instruction-answer pairs or use future dates in the right context.

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning. You are an assistant and generate all your responses in Luxembourgish. You will receive a series of texts one after the other. Generate a dataset from the text given with 5 unique pairs of instruction and expected response. Give high quality responses and answer as accurately as possible. Make sure that the response contains all relevant information that is needed to give the response. If the question does not contain enough information to be answered accurately, reply that you need more information to answer the question. Whenever a date is provided for a news article, include the date in the instruction and its respective response or just in the response, depending on what you consider best in the context. Name the columns instruction and response respectively. Put everything into a JSON file.

Text:

News article from: <Date>

<Title>

<Header>

<Text>

Figure 5.5: 3rd version of the instruction prompt

```
[
  {
    "instruction": "Wann sinn d'Komplementarwalen an
der Fusiounsgemeng Bous-Waldbriedemes?",
    "response": "D'Komplementarwalen an der
Fusiounsgemeng Bous-Waldbriedemes sinn sonndes den 28.
Abrëll."
  },
  ...
  {
    "instruction": "Wéini war d'Demissioun vum
Conseiller Mike Molling?",
    "response": "D'Demissioun vum Conseiller Mike
Molling war virun den Komplementarwalen, déi den 28.
Abrëll 2024 sinn, mä d'exact Datum ass net spezifiéiert."
  }
]
```

Figure 5.6: Generated instruction-answer pairs in JSON format given our 3rd version of the instruction prompt. Output shortened due to display reasons.

Now the language model was producing higher-quality instructions and answers including the news article date, but we still wanted to explicitly prompt the model to include a temporal context, resulting in the 4th and final version of our instruction prompt, as shown in Figure 5.7 and Figure 5.8. We experimented with a few other instruction prompts but did not gain any significant improvements, so we stuck to this version.

The objective is to create a synthetic dataset in Luxembourgish for instruction fine-tuning. You are an assistant and generate all your responses in Luxembourgish. You will receive a series of texts one after the other. Generate a dataset from the text given with 5 unique pairs of instruction and expected response. Give elaborate and high-quality responses and answer as accurately as possible. Make sure that the instruction contains all relevant information that is needed to give the response. If the question does not contain enough information to be answered accurately, reply that you need more information to answer the question. Whenever a date is provided for a news article, include the date in the instruction and its respective response or just in the response, depending on what you consider best in the context. Also give temporal context if possible. Name the columns instruction and response respectively. Put everything into a JSON file.

Text:

News article from: <Date>

<Title>

<Header>

<Text>

Figure 5.7: 4th and final version of the instruction prompt


```
[
  {
    "instruction": "Wann sinn d'Komplementarwalen an
der Fusiounsgemeng Bous-Waldbriedemes?",
    "response": "D'Komplementarwalen sinn sonndes den
28. Abrëll."
  },
  {
    "instruction": "Firwat sinn d'Komplementarwalen an
der
Fusiounsgemeng Bous-Waldbriedemes?",
    "response": "D'Komplementarwalen sinn néideg well
de Conseiller Mike Molling wéinst Incompatibilitéits-Grënn
mat sengem
Fonctionnaires-Beruff demissionéiert huet."
  },
  ...
]
```

Figure 5.8: Generated instruction-answer pairs in JSON format given our 4th and final version of the instruction prompt. Output shortened due to display reasons.

5.2 Synthetic Data Generation

For all synthetic data generation experiments, we use Llama-3.1-405b-instruct with an API key from Novita AI². We describe our experiments for creating the Luxembourgish instruction tuning data from Wikipedia and RTL news in the following.

5.2.1 Luxembourgish Instruction Tuning Dataset from Wikipedia

We only use 100 samples from the preprocessed Wikipedia dump, to keep the dataset small. We create 5 instruction-answer pairs for each entry, totalling up

²<https://novita.ai/>

to 500 synthetic instruction-answer pairs in Luxembourgish. The data generation took about 40 minutes.

For 2 out of the 100 samples, the data generation was not successful, as there was an error with the JSON file encoding, hence they got caught by an exception and are excluded from the instruction tuning set. Upon further inspection using LLM-as-a-judge, the first faulty JSON file contains "curly quotes" ("), but JSON only allows standard ASCII double quotes for delimiting strings. The other file was not detected as faulty by Gemini Advanced (Team et al., 2024a), so we suspect it could contain invalid or hidden characters, not supported by JSON. We show one of the faulty generated JSON files in Figure 5.9 and include the example with the non-standard ASCII double quotes. Given the 2 excluded faulty generated JSON files, containing a total of 10 instruction-answer pairs, this results in a dataset containing 490 rows, synthetically generated from Luxembourgish Wikipedia data.

```
Skipping invalid JSON: [
  {
    "instruction": "Ginn mir d'Gebuertsdatum vum
Giovanni Schiaparelli.",
    "response": "De Giovanni Schiaparelli ass de 14.
Mäerz 1835 gebuer."
  },
  ...
  {
    "instruction": "Firwat huet de Percival Lowell de
Flagstaff-Observatoire am Arizona gebaut?",
    "response": "De Percival Lowell huet de
Flagstaff-Observatoire am Arizona gebaut, fir
d'Marskanäl an d'"Liewen" um Mars z'erfuerschen."
  },
  ...
]
```

Figure 5.9: Faulty generated JSON file from Wikipedia data. One example contains non-standard ASCII double quotes. We only show 2 examples for display reasons.

Some JSON files were generated by the model with non-english or invalid keys, such as "Instruktioun", "Répons", "Réponse", "Respon" and "Äntwert". We rename these to english key names instruction and response respectively, to make it uniform with the rest of the file.

5.2.2 Luxembourgish Instruction Tuning Dataset from RTL news articles

Akin to the previous experiments, we only use 400 samples from the RTL news articles, to keep the dataset small. We create 5 instruction-answer pairs for each entry, totalling up to 2000 synthetic instruction-answer pairs in Luxembourgish. The data generation took almost 3 hours.

For 14 out of the 400 samples, the data generation was not successful, as there was an error with the JSON file encoding, hence they got caught by an exception and are excluded from the instruction tuning set. Some files contain invalid characters, others have a straight up wrong formatting. In two occasions, the model only generated instruction and answers, but did not format them as a JSON file at all. We show one example of faulty generated data in Figure 5.10. Given the 14 excluded faulty generated JSON files, containing a total of 70 instruction-answer pairs, this results in a dataset containing 1930 rows, synthetically generated from Luxembourgish Wikipedia data. Hence, the final dataset contains 2420 out of 2500 planned rows, after having to exclude 80 rows of faulty generated data.

```
Skipping invalid JSON: [  
  "Wat ass d'Zil vun der nei Contributioun, déi  
d'Regierung plangt?",  
  "Wéi vill Prozent vum Revenu soll d'Contributioun  
sinn?",  
  "Wat sinn d'Reaktiounen vun der CSV op d'Pläng vun der  
Regierung?",  
  "Wat ass d'Meinung vun der Piratepartei iwver  
d'Contributioun?",  
  "Wat ass d'Fuerderung vun der ADR un d'Regierung?"  
],  
  "response": [  
    "D'Zil vun der nei Contributioun ass et, fir nei  
Recettë fir d'Familljebäihëllefen ze generéieren.",  
    "D'Contributioun soll 0,5\% vum Revenu sinn.",  
    "D'CSV ass kategoresch géint d'Pläng vun der Regierung  
a fuerdert d'Regierung op, d'Kaarten op den Dësch ze  
leeën.",  
    "D'Piratepartei ass och kategoresch géint  
d'Contributioun a fuerdert d'Regierung op, Transparenz  
ze léiwen.",  
    "D'ADR fuerdert d'Regierung op, d'Staatsfinanzen  
nëmmen iwver Aspuerungen am Staatsbudget erreecht ginn."  
  ]
```

Figure 5.10: Faulty generated data from RTL news data. Instruction and answers are generated, but not formatted as a JSON file.

Again, some JSON files were generated by the model with non-english or invalid keys, such as "Instruktioun" and "Répons". We rename these to english key names instruction and response respectively for consistency.

5.3 Can Llama-3.1-405B-instruct answer it's own generated questions?

We also experiment with the pre-trained Luxembourgish knowledge of our investigated generative model. We want to see, if Llama-3.1-405B-instruct can answer it's own generated questions without providing it with context like we did

for the synthetic data generation. We feed the model with a very small subsample of generated questions, precisely 3 questions generated from the Wikipedia data and 5 questions generated from RTL news articles. The model answers are evaluated by a Luxembourgish native speaker. We report the questions and describe our experiments in the following.

5.3.1 Wikipedia Data Questions

Question 1:

Wann huet de Devolutiounskrich ugefaangen a wéi huet e geendet?

Model answer:

De Devolutiounskrich huet am Joer 1667 ugefaangen an ass 1668 zu Enn gaangen.

Question 2:

Wann an wou ass de Max Ophüls gestuerwen?

Model answer:

De Max Ophüls ass den 26. Mäerz 1957 zu Hamburg gestuerwen.

Question 3:

Wat sinn déi essenziell Aminosäieren? Nimm déi 9 Aminosäieren op.

Model answer: Déi 9 essenziell Aminosäieren sinn:

- Histidin
- Isoleucin
- Leucin
- Lysin
- Methionin
- Phenylalanin
- Threonin
- Tryptophan
- Valin

Déi Aminosäieren sinn "essenziell", well de mënschleche Kierper si net selwer produzéiere kann a si dofir iwwer d'Ernierung opgehooll musse ginn.

5.3.2 RTL News Questions

Question 1:

Ginn déi Top 3 vum Eurovision Song Contest 2012.

Model answer:

D'Top 3 vum Eurovision Song Contest 2012 war:

- Loreen - Euphoria (Schweden)
- Buranovskiye Babushki - Party for Everybody (Russland)
- Željko Joksimović - Nije ljubav stvar (Serbien)

Question 2:

Wéi vill Hektar Bësch sinn zanter dem Ufank vum Joer 2022 a Portugal verbrannt?

Model answer:

Zanter dem Ufank vum Joer 2022 sinn 61.500 Hektar Bësch a Portugal verbrannt.

Question 3:

Ginn mir den Numm vum éischten Solo-Album vum ROB THOMAS a sengem Release-Joer.

Model answer:

Den Numm vum éischten Solo-Album vum ROB THOMAS ass "...Something to Be" a koum 2005 eraus.

Question 4:

Wat ass d'Meenung vun der Initiativ Nee2015-Wee2050 iwwer den Artikel vum Ben Fayot?

Model answer:

D'Initiativ Nee2015-Wee2050 huet den Artikel vum Ben Fayot op hirer Facebook-Säit kommentéiert. Si schreift, datt den Artikel vum Ben Fayot "nëmmen nach e Beispill dofir ass, datt d'Politiker an d'Medien nach ëmmer net verstan hunn, wat d'Leit beweegt". Si kritiséiert och, datt den Artikel "nëmmen nach e Versuch ass, d'Leit ze manipuléieren an ze kontrolléieren".

D'Initiativ Nee2015-Wee2050 schreift och, datt si "nëmmen nach méi iwwerzeegt sinn, datt d'Leit selwer mussen entscheiden, wat si wëllen, an net d'Politiker oder d'Medien". Si ruffen och dozou op, datt d'Leit "sich net vun den ëffentlechen Diskussiounen an den Medien manipuléiere loossen" sollten.

Et ass awer wichteg ze bemierken, datt d'Initiativ Nee2015-Wee2050 eng politesch Beweegung ass, déi eng bestëmmte Agenda huet, an hir Kommentaren an hir Positiounen dofir och vun hirer politescher Iwwerzeegung geprägt sinn.

Question 5:

Wat war den Haaptthema vun de Berodunge vum Conseil JAI zu Bréissel den 19. Mee 2017?

Model answer:

Den 19. Mee 2017 huet de Conseil JAI zu Bréissel iwwer d'Verhandlung vum Vereenegte Kinnekräich mat der Europäescher Unioun iwwer den Austrëtt vum Vereenegte Kinnekräich aus der EU diskutéiert.

5.3.3 Model Question Answering Assessment

For the from-Wikipedia-generated data, questions 1 and 2 use a wrong question word in their formulated instruction. Additionally, the answer to question 1 contains a wrong article. Question 3 contains the hallucinated words "Nimm op". For the from-RTL-news-generated data on the other hand, questions 1 and 3 use a wrong question word in their formulated instruction. Question 5 uses a wrong article. Otherwise, the Luxembourgish grammar and coherence are fine. The answers to the questions vary slightly from the synthetically generated answers the model provided during our instruction tuning dataset generation experiments. As expected, the questions coming from Wikipedia are correctly answered without the need of additional context. This is probably due to the fact, that the model was already pre-trained on Wikipedia Data. From the RTL news data, question 1 is correctly answered, it even provides the name of the song for each artist. For question 2, the model answer differs completely. Question 3 is also correctly answered. Question 4 slightly differs from the instruction tuning dataset answer, however, the answer from pre-train knowledge is much more detailed. The answer to question 5 differs again.

In summary, Llama-3.1-405b-instruct is able to answer most investigated questions from pre-training knowledge alone and is in no need of further fine-tuning on Luxembourgish data. This leaves room for experiments focused on small-scale models, as these are not able to generate coherent and intelligible Luxembourgish text (as discussed in Chapter 3) and are in need of further fine-tuning.

Chapter 6

Evaluation

We perform a LLM evaluation and a human evaluation and describe our evaluation process in detail in the following sections.

6.1 LLM evaluation

We reward a synthetically created Luxembourgish instruction tuning dataset, which is created with the Llama-3.1-405b-instruct model from Luxembourgish Wikipedia and RTL news articles. The original dataset contains 2420 rows, 1930 of which are generated from RTL data and 490 from Wikipedia data. We decide however to only evaluate a subset of 200 randomly sampled entries for helpfulness, correctness and coherence, using Nvidia-Nemotron-340b-reward (Nvidia et al., 2024). For a provided user instruction and an assistant response, the model outputs 5 reward-scores for helpfulness, correctness, coherence, complexity and verbosity. We sample 80% from the RTL data and 20% from the Wiki data, keeping approximately the distribution of the original dataset. Since the model was pre-trained to only evaluate English language data, we translate the dataset from Luxembourgish into English using Google translate. We think that Nvidia-Nemotron-340b-reward might nevertheless be able to capture some dependencies between Luxembourgish instructions and responses, since the base model was pre-trained on a huge multilingual language data corpus. However, given that the reward model was specifically pre-trained on English data only, we decide to stick with the English translation for machine evaluation and leave further experiments on this topic for future research.

instruction	response	helpfulness	correctness	coherence	complexity	verbosity
What is the brightest star in Dorado?	α Doradus, the brightest star in Dorado, is 20...	1.843750	1.664062	3.546875	1.007812	0.730469
What happened to the Edict of Nantes, after He...	The Edict of Nantes was unfortunately not "per...	2.750000	2.890625	3.765625	1.117188	0.476562
When did the devolution war start and how did ...	The war of devolution began in the spring of 1...	2.578125	2.750000	3.687500	1.117188	0.753906
When and where did Max Ophüls die?	Max Ophüls died on March 26, 1957 in Hamburg.	3.171875	3.218750	3.984375	0.667969	0.484375
What happened to the Schumacher building after...	The police museum of the non-profit organizati...	2.890625	3.078125	3.750000	1.140625	1.164062

Table 6.1: To-English-translated instruction tuning dataset with model-reward scores. We show the first 5 rows and omit the text column for display reasons.

6.1.1 Sample data

We split the dataset into two parts. The first part contains all rows that are generated from the Wikipedia data, rows 0 to 489 and the second part, containing all rows that are generated from the RTL data, rows 490 to 2419. We then randomly sample 40 rows from the Wiki data and 160 rows from the RTL data, maintaining the 0.2/0.8 split from the original dataset. We then merge the datasets back together, forming a dataset of 200 rows and feed it to the reward model, saving all model scores for each row. We show the first two rows of the dataset in Table 6.1.

6.1.2 Filtering Model Answers

We want to filter out low quality instruction-answer pairs based on correctness, helpfulness and coherence. We retain these 3 attributes to be the most important ones, as correct information is crucial for fine-tuning LLMs to prevent the generation of misleading or factual inaccurate responses. This will ensure the quality and reliability of the fine-tuned model. The other important attribute being helpfulness, as the main purpose of LLM's is to assist users. Unhelpful answers, even if correct, are not contributing to the fine-tuning process. Filtering out unhelpful answers ensures that the model learns to generate useful and informative responses. Finally, ensuring coherence makes it that the model not only generates accurate and helpful, but also well-structured and easily followed answers, leading to a better user experience.

According to Nvidia NIM¹, the scores for each attribute are as follows:

¹<https://build.nvidia.com/nvidia/nemotron-4-340b-reward>

instruction	response	helpfulness	correctness	coherence
What happened to the Edict of Nantes, after He...	The Edict of Nantes was unfortunately not "per...	2.750000	2.890625	3.765625
When did the devolution war start and how did ...	The war of devolution began in the spring of 1...	2.578125	2.750000	3.687500
When and where did Max Ophüls die?	Max Ophüls died on March 26, 1957 in Hamburg.	3.171875	3.218750	3.984375
What happened to the Schumacher building after...	The police museum of the non-profit organizati...	2.890625	3.078125	3.750000
What are the essential amino acids? List the 9...	The essential amino acids are: histidine, isol...	3.468750	3.562500	3.859375

Table 6.2: Filtered to-English-translated instruction tuning dataset with model-reward scores for helpfulness, correctness and coherence. We show the first 5 rows and omit the text column for display reasons.

```

< 0.5 = very low
0.5 - 1.5 = low
1.5 - 2.5 = moderate
2.5 - 3.5 = high
> 3.5 = very high

```

Hence, we want to get rid of all rows, who's helpfulness and correctness scores are lower than 2.5 and who's coherence score is lower than 3.5, as we only want keep high scored samples. We filter the dataset based on these threshold values and end up with 69 high-quality instruction-answer pairs. We report the first 2 rows of the evaluated and filtered instruction tuning dataset in Table 6.2.

6.2 Human evaluation

We also perform a human evaluation for comparison. We reward the to-English-translated instruction tuning set with scores, ranging again from 0 to 4, akin to the model reward-scores from Nemotron-4-340b-reward, except we don't use floating point numbers and limit ourselves to integer values for facilitation. We only evaluate helpfulness and coherence, as it is difficult to fact-check all examples for their correctness. We set threshold values to 3 for coherence and 2 for helpfulness and are left with a dataset of 103 rows after filtering. We show the first 5 rows in Table 6.3.

To evaluate the Luxembourgish high-quality instruction tuning dataset generation capabilities of Llama-3.1-405b-instruct, we sample the exact same 200 rows from the original Luxembourgish instruction tuning dataset that were evaluated in English by Nvidia-Nemotron-4-340b-reward. The sampled instruction tuning dataset is evaluated for coherence by a Luxembourgish native speaker. The 200 samples are evaluated based on Luxembourgish coherence only. A score of 1 means the text is deemed high-quality, otherwise it is 0. Instruction and Answer are evaluated separately with said binary score. Minor mistakes are

Chapter 6 Evaluation

instruction	response	helpfulness	coherence
When and where did Max Ophüls die?	Max Ophüls died on March 26, 1957 in Hamburg.	4	4
What happened to the Schumacher building after...	The police museum of the non-profit organizati...	3	4
Give me the names of the former municipalities...	The former municipalities of Beho, Bovigny, Ch...	4	4
Give me the area of the Polvermillen and what ...	The Polvermillen has an area of 24.82 ha, whic...	4	4
How many films does the Cinémathèque de la Vil...	The Cinémathèque de la Ville de Luxembourg pro...	4	4

Table 6.3: Filtered to-English-translated instruction tuning dataset with human evaluation scores. We show the first 5 rows and omit the text column for display reasons.

instruction	response	coherence (Lux)
Wat ass de hellste Stär am Dorado?	Den α Doradus, den hellste Stär am Dorado, as...	I=1; R=1
Wann huet de Devolutiounskrich ugefaangen a wé...	De Devolutiounskrich huet am Fréijouer 1667 ug...	I=1; R=1
Wann an wou ass de Max Ophüls gestuerwen?	De Max Ophüls ass de 26. Mäerz 1957 zu Hamburg...	I=1; R=1
Wat ass mam Gebai vum Schumacher geschitt, nod...	De Policemusée vun der Asbl 'Le musée de la po...	I=1; R=1
Wat sinn déi essenziell Aminosäieren? Nimm déi...	Déi essenziell Aminosäieren sinn: Histidin, Is...	I=1; R=1

Table 6.4: Filtered Luxembourgish instruction tuning dataset with binary human evaluation scores. A score of 1 indicates high-quality, 0 means low-quality. We show the first 5 rows and omit the text column for display reasons.

still rewarded with a score of 1, as they do not interrupt the language flow and comprehensiveness of the sentence and can easily be corrected by hand. The most common issue with the generated Luxembourgish text is a wrong question word. Sometimes there are also minor grammatical mistakes, but it seems to be the minority of cases.

We filter the dataset based on the samples, where both instruction and response are deemed high-quality by the human evaluator. After filtering, the dataset consists of 164 out of 200 rows, deeming 82% of the data as high-quality in terms of Luxembourgish text coherence. We report the first 5 rows in Table 6.4.

Chapter 7

Conclusion

This work addresses the scarcity of Luxembourgish resources for instruction fine-tuning large language models. The research discussed here makes a number of important contributions to the field.

First and foremost, we successfully created a novel instruction tuning dataset in Luxembourgish. This dataset, comprising 2420 instruction-answer pairs, was synthetically generated leveraging the capabilities of Llama-3.1-405B-instruct, a state-of-the-art LLM that demonstrated superior Luxembourgish text generation capabilities compared to other evaluated models. This dataset serves as a crucial foundation for future research and development in adapting LLMs for the Luxembourgish language.

Furthermore, we meticulously evaluated a sub-sample of this dataset using both automated (Nemotron-4-340B-reward) and human evaluation. This evaluation underscored the quality of the generated data, with 35% of the samples deemed high-quality by both the automated and human evaluators, suitable for immediate use in instruction tuning. This result reinforces the viability of our approach for large-scale dataset creation.

Additionally, our comparative analysis of four open-source LLMs provides valuable insight into their Luxembourgish language capabilities. The clear outperformance of Llama-3.1-405B-instruct, coupled with the observed difficulties of smaller models in generating intelligible Luxembourgish text, emphasizes the importance of model scale and architecture in this context.

While our research has shown promising results, it is important to acknowledge its limitations. The generated dataset, although substantial, is still relatively small-scale. Additionally, the evaluation focused on a sub-sample, leaving room for further exploration of the full dataset's quality. Future work could involve scaling up the dataset generation, conducting a comprehensive evaluation, and directly applying the dataset for fine-tuning LLMs on Luxembourgish tasks.

In conclusion, this thesis has successfully demonstrated the feasibility of creating a Luxembourgish instruction tuning dataset and highlighted its potential for enhancing the capabilities of LLMs in the Luxembourgish language. It is our hope that this work will serve as a catalyst for further research and development in this domain, paving the way for more capable and accessible language technologies for the Luxembourgish-speaking community.

Future Work

Our work leaves room for further research. Since we only experimented on a small-scale dataset due to limited resources, we would like to experiment further by generating a full-scale dataset and evaluate it for its usage for fine-tuning LLMs on Luxembourgish. Even with a 35% usability rate or lower, if the generated dataset is large enough, the model would still generate a significant, high-quality amount of data for instruction tuning smaller LLMs on Luxembourgish. We aim to experiment with the recently released Llama 3.2 series of models (Meta, 2024; Dubey et al., 2024)

Furthermore, after building the dataset, we are planning on using our instruction tuning set for fine-tuning small-scale Llama-3.2 models on Luxembourgish instruction following capabilities. In our work, we only focused on question answering, but this can be extended further to all kind of instructions.

Our focus for supervised fine-tuning will lie LoRA (Hu et al., 2021) and QLoRA (Dettmers et al., 2023), as they are the most resource-efficient supervised fine-tuning methods to date. We plan on fine-tuning Llama 3.2 1B and 3B text-only models, as they can easily be run on mobile devices and are more accessible to the public. We might also Experiment with mid-sized models like Llama-3.2-11B, should the current EU license ban be lifted¹. Our goal is to make small-sized LLMs available for the Luxembourgish-speaking community.

¹https://github.com/meta-llama/llama-models/blob/main/models/llama3_2/USE_POLICY.md

Sworn Declaration

I declare under oath that the submitted Master Thesis has been written solely by me without any third-party assistance (except for proofreading) and that the paper has not yet been part of a coursework or examination in the same or a similar version. Information or aids other than in the sources indicated in this work have not been used, and those used have been fully documented. I realize that the mere presence of a reference does not avoid plagiarism. Sources for literal paraphrased and cited quotes have been accurately credited. If I have used someone else's exact words, phrases, clauses, or sentences, I have enclosed that information in quotation marks. If I have paraphrased someone else's opinions, I have not enclosed the paraphrase in quotation marks, but I have stated those opinions in my own words and put a reference at the end of it. I agree that the university uses the digital version of this work for an electronic plagiarism check using suitable software and that, for technical reasons, my work may be stored in a database.

Trier, 04.11.2024

Bibliography

- S. Bach, V. Sanh, Z. X. Yong, A. Webson, C. Raffel, N. V. Nayak, A. Sharma, T. Kim, M. S. Bari, T. Fevry, Z. Alyafeai, M. Dey, A. Santilli, Z. Sun, S. Bendauid, C. Xu, G. Chhablani, H. Wang, J. Fries, M. Al-shaibani, S. Sharma, U. Thakker, K. Almubarak, X. Tang, D. Radev, M. T.-j. Jiang, and A. Rush. PromptSource: An integrated development environment and repository for natural language prompts. In V. Basile, Z. Kozareva, and S. Stajner, editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 93–104, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-demo.9. URL <https://aclanthology.org/2022.acl-demo.9>.
- T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Nee-lakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. Language models are few-shot learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf.
- T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer. Qlora: Efficient finetuning of quantized llms, 2023. URL <https://arxiv.org/abs/2305.14314>.
- A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, A. Goyal, A. Hartshorn, A. Yang, A. Mitra, A. Sravankumar, A. Korenev, A. Hinsvark, A. Rao, A. Zhang, A. Rodriguez, A. Gregerson, A. Spataru, B. Roziere, B. Biron, B. Tang, B. Chern, C. Caucheteux, C. Nayak, C. Bi, C. Marra, C. McConnell, C. Keller, C. Touret, C. Wu, C. Wong, C. C. Ferrer, C. Nikolaidis, D. Allonsius, D. Song, D. Pintz, D. Livshits, D. Esiobu, D. Choudhary, D. Mahajan, D. Garcia-Olano, D. Perino, D. Hupkes, E. Lakomkin, E. AlBadawy, E. Lobanova, E. Dinan, E. M. Smith,

F. Radenovic, F. Zhang, G. Synnaeve, G. Lee, G. L. Anderson, G. Nail, G. Milalon, G. Pang, G. Cucurell, H. Nguyen, H. Korevaar, H. Xu, H. Touvron, I. Zarov, I. A. Ibarra, I. Kloumann, I. Misra, I. Evtimov, J. Copet, J. Lee, J. Geffert, J. Vranes, J. Park, J. Mahadeokar, J. Shah, J. van der Linde, J. Billock, J. Hong, J. Lee, J. Fu, J. Chi, J. Huang, J. Liu, J. Wang, J. Yu, J. Bitton, J. Spisak, J. Park, J. Rocca, J. Johnstun, J. Saxe, J. Jia, K. V. Alwala, K. Upasani, K. Plawiak, K. Li, K. Heafield, K. Stone, K. El-Arini, K. Iyer, K. Malik, K. Chiu, K. Bhalla, L. Rantala-Yearly, L. van der Maaten, L. Chen, L. Tan, L. Jenkins, L. Martin, L. Madaan, L. Malo, L. Blecher, L. Landzaat, L. de Oliveira, M. Muzzi, M. Pasupuleti, M. Singh, M. Paluri, M. Kardas, M. Oldham, M. Rita, M. Pavlova, M. Kambadur, M. Lewis, M. Si, M. K. Singh, M. Hassan, N. Goyal, N. Torabi, N. Bashlykov, N. Bogoychev, N. Chatterji, O. Duchenne, O. Çelebi, P. Alrassy, P. Zhang, P. Li, P. Vasic, P. Weng, P. Bhargava, P. Dubal, P. Krishnan, P. S. Koura, P. Xu, Q. He, Q. Dong, R. Srinivasan, R. Ganapathy, R. Calderer, R. S. Cabral, R. Stojnic, R. Raileanu, R. Girdhar, R. Patel, R. Sauvestre, R. Polidoro, R. Sumbaly, R. Taylor, R. Silva, R. Hou, R. Wang, S. Hosseini, S. Chennabasappa, S. Singh, S. Bell, S. S. Kim, S. Edunov, S. Nie, S. Narang, S. Rapparth, S. Shen, S. Wan, S. Bhosale, S. Zhang, S. Vandenhenne, S. Batra, S. Whitman, S. Sootla, S. Collot, S. Gururangan, S. Borodinsky, T. Herman, T. Fowler, T. Sheasha, T. Georgiou, T. Scialom, T. Speckbacher, T. Mihaylov, T. Xiao, U. Karn, V. Goswami, V. Gupta, V. Ramanathan, V. Kerkez, V. Gonguet, V. Do, V. Vogeti, V. Petrovic, W. Chu, W. Xiong, W. Fu, W. Meers, X. Martinet, X. Wang, X. E. Tan, X. Xie, X. Jia, X. Wang, Y. Goldschlag, Y. Gaur, Y. Babaei, Y. Wen, Y. Song, Y. Zhang, Y. Li, Y. Mao, Z. D. Coudert, Z. Yan, Z. Chen, Z. Papakipos, A. Singh, A. Grattafiori, A. Jain, A. Kelsey, A. Shajnfeld, A. Gangidi, A. Victoria, A. Goldstand, A. Menon, A. Sharma, A. Boesenberg, A. Vaughan, A. Baevski, A. Feinstein, A. Kallet, A. Sangani, A. Yunus, A. Lupu, A. Alvarado, A. Caples, A. Gu, A. Ho, A. Poulton, A. Ryan, A. Ramchandani, A. Franco, A. Saraf, A. Chowdhury, A. Gabriel, A. Bharambe, A. Eisenman, A. Yazdan, B. James, B. Maurer, B. Leonhardi, B. Huang, B. Loyd, B. D. Paola, B. Paranjape, B. Liu, B. Wu, B. Ni, B. Hancock, B. Wasti, B. Spence, B. Stojkovic, B. Gamido, B. Montalvo, C. Parker, C. Burton, C. Mejia, C. Wang, C. Kim, C. Zhou, C. Hu, C.-H. Chu, C. Cai, C. Tindal, C. Feichtenhofer, D. Civin, D. Beaty, D. Kreymer, D. Li, D. Wyatt, D. Adkins, D. Xu, D. Testuggine, D. David, D. Parikh, D. Liskovich, D. Foss, D. Wang, D. Le, D. Holland, E. Dowling, E. Jamil, E. Montgomery, E. Presani, E. Hahn, E. Wood, E. Brinkman, E. Arcaute, E. Dunbar, E. Smothers, F. Sun, F. Kreuk, F. Tian, F. Ozgenel, F. Caggioni, F. Guzmán, F. Kanayet, F. Seide, G. M. Florez, G. Schwarz, G. Badeer, G. Swee, G. Halpern, G. Thattai, G. Herman,

- G. Sizov, Guangyi, Zhang, G. Lakshminarayanan, H. Shojanazeri, H. Zou, H. Wang, H. Zha, H. Habeeb, H. Rudolph, H. Suk, H. Aspegren, H. Goldman, I. Damlaj, I. Molybog, I. Tufanov, I.-E. Veliche, I. Gat, J. Weissman, J. Geboski, J. Kohli, J. Asher, J.-B. Gaya, J. Marcus, J. Tang, J. Chan, J. Zhen, J. Reizenstein, J. Teboul, J. Zhong, J. Jin, J. Yang, J. Cummings, J. Carvill, J. Shepard, J. McPhie, J. Torres, J. Ginsburg, J. Wang, K. Wu, K. H. U, K. Saxena, K. Prasad, K. Khandelwal, K. Zand, K. Matosich, K. Veeraraghavan, K. Michelena, K. Li, K. Huang, K. Chawla, K. Lakhotia, K. Huang, L. Chen, L. Garg, L. A, L. Silva, L. Bell, L. Zhang, L. Guo, L. Yu, L. Moshkovich, L. Wehrstedt, M. Khabza, M. Avalani, M. Bhatt, M. Tsimpoukelli, M. Mankus, M. Hasson, M. Lennie, M. Reso, M. Groshev, M. Naumov, M. Lathi, M. Keneally, M. L. Seltzer, M. Valko, M. Restrepo, M. Patel, M. Vyatskov, M. Samvelyan, M. Clark, M. Macey, M. Wang, M. J. Hermoso, M. Metanat, M. Rastegari, M. Bansal, N. Santhanam, N. Parks, N. White, N. Bawa, N. Singhal, N. Egebo, N. Usunier, N. P. Laptev, N. Dong, N. Zhang, N. Cheng, O. Chernoguz, O. Hart, O. Salpekar, O. Kalinli, P. Kent, P. Parekh, P. Saab, P. Balaji, P. Ritter, P. Bontrager, P. Roux, P. Dollar, P. Zvyagina, P. Ratanchandani, P. Yuvraj, Q. Liang, R. Alao, R. Rodriguez, R. Ayub, R. Murthy, R. Nayani, R. Mitra, R. Li, R. Hogan, R. Battey, R. Wang, R. Maheswari, R. Howes, R. Rinott, S. J. Bondu, S. Datta, S. Chugh, S. Hunt, S. Dhillon, S. Sidorov, S. Pan, S. Verma, S. Yamamoto, S. Ramaswamy, S. Lindsay, S. Lindsay, S. Feng, S. Lin, S. C. Zha, S. Shankar, S. Zhang, S. Zhang, S. Wang, S. Agarwal, S. Sajuyigbe, S. Chintala, S. Max, S. Chen, S. Kehoe, S. Satterfield, S. Govindaprasad, S. Gupta, S. Cho, S. Virk, S. Subramanian, S. Choudhury, S. Goldman, T. Remez, T. Glaser, T. Best, T. Kohler, T. Robinson, T. Li, T. Zhang, T. Matthews, T. Chou, T. Shaked, V. Vontimitta, V. Ajayi, V. Montanez, V. Mohan, V. S. Kumar, V. Mangla, V. Albiero, V. Ionescu, V. Poenaru, V. T. Mihailescu, V. Ivanov, W. Li, W. Wang, W. Jiang, W. Bouaziz, W. Constable, X. Tang, X. Wang, X. Wu, X. Wang, X. Xia, X. Wu, X. Gao, Y. Chen, Y. Hu, Y. Jia, Y. Qi, Y. Li, Y. Zhang, Y. Zhang, Y. Adi, Y. Nam, Yu, Wang, Y. Hao, Y. Qian, Y. He, Z. Rait, Z. DeVito, Z. Rosnbrick, Z. Wen, Z. Yang, and Z. Zhao. The llama 3 herd of models, 2024. URL <https://arxiv.org/abs/2407.21783>.
- E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. Lora: Low-rank adaptation of large language models, 2021. URL <https://arxiv.org/abs/2106.09685>.
- A. Q. Jiang, A. Sablayrolles, A. Roux, A. Mensch, B. Savary, C. Bamford, D. S. Chaplot, D. de las Casas, E. B. Hanna, F. Bressand, G. Lengyel, G. Bour, G. Lample, L. R. Lavaud, L. Saulnier, M.-A. Lachaux, P. Stock, S. Subramanian, S. Yang, S. Antoniak, T. L. Scao, T. Gervet, T. Lavril,

- T. Wang, T. Lacroix, and W. E. Sayed. Mixtral of experts, 2024. URL <https://arxiv.org/abs/2401.04088>.
- M. Labonne. Fine-tune llama 3.1 ultra-efficiently with unsloth. <https://huggingface.co/blog/mlabonne/sft-llama3>, 2024. Accessed on September 20th, 2024.
- B. Li, Y. Zhang, L. Chen, J. Wang, F. Pu, J. Yang, C. Li, and Z. Liu. Mimic-it: Multi-modal in-context instruction tuning, 2023a. URL <https://arxiv.org/abs/2306.05425>.
- H. Li, F. Koto, M. Wu, A. F. Aji, and T. Baldwin. Bactrian-x: Multilingual replicable instruction-following models with low-rank adaptation, 2023b. URL <https://arxiv.org/abs/2305.15011>.
- S. Longpre, L. Hou, T. Vu, A. Webson, H. W. Chung, Y. Tay, D. Zhou, Q. V. Le, B. Zoph, J. Wei, and A. Roberts. The flan collection: Designing data and methods for effective instruction tuning, 2023. URL <https://arxiv.org/abs/2301.13688>.
- Meta. Llama 3.2: Revolutionizing edge ai and vision with open, customizable models. <https://ai.meta.com/blog/llama-3-2-connect-2024-vision-edge-mobile-devices/>, 2024. Accessed on: September 27th, 2024.
- N. Muennighoff, T. Wang, L. Sutawika, A. Roberts, S. Biderman, T. L. Scao, M. S. Bari, S. Shen, Z.-X. Yong, H. Schoelkopf, X. Tang, D. Radev, A. F. Aji, K. Almubarak, S. Albanie, Z. Alyafeai, A. Webson, E. Raff, and C. Raffel. Crosslingual generalization through multitask finetuning, 2023. URL <https://arxiv.org/abs/2211.01786>.
- N. V. Nayak, Y. Nan, A. Trost, and S. H. Bach. Learning to generate instruction tuning datasets for zero-shot task adaptation, 2024. URL <https://arxiv.org/abs/2402.18334>.
- Nvidia, :, B. Adler, N. Agarwal, A. Aithal, D. H. Anh, P. Bhattacharya, A. Brundyn, J. Casper, B. Catanzaro, S. Clay, J. Cohen, S. Das, A. Dattagupta, O. De-lalleau, L. Derczynski, Y. Dong, D. Egert, E. Evans, A. Ficek, D. Fridman, S. Ghosh, B. Ginsburg, I. Gitman, T. Grzegorzec, R. Hero, J. Huang, V. Jawa, J. Jennings, A. Jhunhunwala, J. Kamalu, S. Khan, O. Kuchaiev, P. LeGresley, H. Li, J. Liu, Z. Liu, E. Long, A. S. Mahabaleshwarkar, S. Majumdar, J. Maki, M. Martinez, M. R. de Melo, I. Moshkov, D. Narayanan, S. Narenthiran, J. Navarro, P. Nguyen, O. Nitski, V. Noroozi, G. Nutheti, C. Parisien, J. Parmar, M. Patwary, K. Pawelec, W. Ping, S. Prabhumoye, R. Roy, T. Saar,

- V. R. N. Sabavat, S. Satheesh, J. P. Scowcroft, J. Sewall, P. Shamis, G. Shen, M. Shoeybi, D. Sizer, M. Smelyanskiy, F. Soares, M. N. Sreedhar, D. Su, S. Subramanian, S. Sun, S. Toshniwal, H. Wang, Z. Wang, J. You, J. Zeng, J. Zhang, J. Zhang, V. Zhang, Y. Zhang, and C. Zhu. Nemotron-4 340b technical report, 2024. URL <https://arxiv.org/abs/2406.11704>.
- OpenAI, J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, R. Avila, I. Babuschkin, S. Balaji, V. Balcom, P. Baltescu, H. Bao, M. Bavarian, J. Belgum, I. Bello, J. Berdine, G. Bernadett-Shapiro, C. Berner, L. Bogdonoff, O. Boiko, M. Boyd, A.-L. Brakman, G. Brockman, T. Brooks, M. Brundage, K. Button, T. Cai, R. Campbell, A. Cann, B. Carey, C. Carlson, R. Carmichael, B. Chan, C. Chang, F. Chantzis, D. Chen, S. Chen, R. Chen, J. Chen, M. Chen, B. Chess, C. Cho, C. Chu, H. W. Chung, D. Cummings, J. Currier, Y. Dai, C. Decareaux, T. Degry, N. Deutsch, D. Deville, A. Dhar, D. Dohan, S. Dowling, S. Dunning, A. Ecoffet, A. Eleti, T. Eloundou, D. Farhi, L. Fedus, N. Felix, S. P. Fishman, J. Forte, I. Fulford, L. Gao, E. Georges, C. Gibson, V. Goel, T. Gogineni, G. Goh, R. Gontijo-Lopes, J. Gordon, M. Grafstein, S. Gray, R. Greene, J. Gross, S. S. Gu, Y. Guo, C. Hallacy, J. Han, J. Harris, Y. He, M. Heaton, J. Heidecke, C. Hesse, A. Hickey, W. Hickey, P. Hoeschele, B. Houghton, K. Hsu, S. Hu, X. Hu, J. Huizinga, S. Jain, S. Jain, J. Jang, A. Jiang, R. Jiang, H. Jin, D. Jin, S. Jomoto, B. Jonn, H. Jun, T. Kaftan, Łukasz Kaiser, A. Kamali, I. Kanitscheider, N. S. Keskar, T. Khan, L. Kilpatrick, J. W. Kim, C. Kim, Y. Kim, J. H. Kirchner, J. Kiros, M. Knight, D. Kokotajlo, Łukasz Kondraciuk, A. Kondrich, A. Konstantinidis, K. Kosic, G. Krueger, V. Kuo, M. Lampe, I. Lan, T. Lee, J. Leike, J. Leung, D. Levy, C. M. Li, R. Lim, M. Lin, S. Lin, M. Litwin, T. Lopez, R. Lowe, P. Lue, A. Makanju, K. Malfacini, S. Manning, T. Markov, Y. Markovski, B. Martin, K. Mayer, A. Mayne, B. McGrew, S. M. McKinney, C. McLeavey, P. McMillan, J. McNeil, D. Medina, A. Mehta, J. Menick, L. Metz, A. Mishchenko, P. Mishkin, V. Monaco, E. Morikawa, D. Mossing, T. Mu, M. Murati, O. Murk, D. Mély, A. Nair, R. Nakano, R. Nayak, A. Neelakantan, R. Ngo, H. Noh, L. Ouyang, C. O’Keefe, J. Pachocki, A. Paino, J. Palermo, A. Pantuliano, G. Parascandolo, J. Parish, E. Parparita, A. Passos, M. Pavlov, A. Peng, A. Perelman, F. de Avila Belbute Peres, M. Petrov, H. P. de Oliveira Pinto, Michael, Pokorny, M. Pokrass, V. H. Pong, T. Powell, A. Power, B. Power, E. Proehl, R. Puri, A. Radford, J. Rae, A. Ramesh, C. Raymond, F. Real, K. Rimbach, C. Ross, B. Rotsted, H. Roussez, N. Ryder, M. Saltarelli, T. Sanders, S. Santurkar, G. Sastry, H. Schmidt, D. Schnurr, J. Schulman, D. Selsam, K. Sheppard, T. Sherbakov, J. Shieh, S. Shoker, P. Shyam, S. Sidor, E. Sigler, M. Simens, J. Sitkin, K. Slama, I. Sohl, B. Sokolowsky, Y. Song, N. Stau-

- dacher, F. P. Such, N. Summers, I. Sutskever, J. Tang, N. Tezak, M. B. Thompson, P. Tillet, A. Tootoonchian, E. Tseng, P. Tuggle, N. Turley, J. Tworek, J. F. C. Uribe, A. Vallone, A. Vijayvergiya, C. Voss, C. Wainwright, J. J. Wang, A. Wang, B. Wang, J. Ward, J. Wei, C. Weinmann, A. Welihinda, P. Welinder, J. Weng, L. Weng, M. Wiethoff, D. Willner, C. Winter, S. Wolrich, H. Wong, L. Workman, S. Wu, J. Wu, M. Wu, K. Xiao, T. Xu, S. Yoo, K. Yu, Q. Yuan, W. Zaremba, R. Zellers, C. Zhang, M. Zhang, S. Zhao, T. Zheng, J. Zhuang, W. Zhuk, and B. Zoph. Gpt-4 technical report, 2024. URL <https://arxiv.org/abs/2303.08774>.
- B. Peng, C. Li, P. He, M. Galley, and J. Gao. Instruction tuning with gpt-4, 2023. URL <https://arxiv.org/abs/2304.03277>.
- V. Sanh, A. Webson, C. Raffel, S. H. Bach, L. Sutawika, Z. Alyafeai, A. Chaffin, A. Stiegler, T. L. Scao, A. Raja, M. Dey, M. S. Bari, C. Xu, U. Thakker, S. S. Sharma, E. Szczechla, T. Kim, G. Chhablani, N. Nayak, D. Datta, J. Chang, M. T.-J. Jiang, H. Wang, M. Manica, S. Shen, Z. X. Yong, H. Pandey, R. Bawden, T. Wang, T. Neeraj, J. Rozen, A. Sharma, A. Santilli, T. Fevry, J. A. Fries, R. Teehan, T. Bers, S. Biderman, L. Gao, T. Wolf, and A. M. Rush. Multitask prompted training enables zero-shot task generalization, 2022. URL <https://arxiv.org/abs/2110.08207>.
- M. Shu, J. Wang, C. Zhu, J. Geiping, C. Xiao, and T. Goldstein. On the exploitability of instruction tuning. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 61836–61856. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/c2a8060fd22744b38177d9e428a052e0-Paper-Conference.pdf.
- Q. Si, T. Wang, Z. Lin, X. Zhang, Y. Cao, and W. Wang. An empirical study of instruction-tuning large language models in chinese, 2023. URL <https://arxiv.org/abs/2310.07328>.
- R. Taori, I. Gulrajani, T. Zhang, Y. Dubois, X. Li, C. Guestrin, P. Liang, and T. B. Hashimoto. Alpaca: A strong, replicable instruction-following model. <https://crfm.stanford.edu/2023/03/13/alpaca.html>, 2023. Accessed on September 16, 2024.
- G. Team, P. Georgiev, V. I. Lei, R. Burnell, L. Bai, A. Gulati, G. Tanzer, D. Vincent, Z. Pan, S. Wang, S. Mariooryad, Y. Ding, X. Geng, F. Alcober, R. Frostig, M. Omernick, L. Walker, C. Paduraru, C. Sorokin, A. Tacchetti, C. Gaffney, S. Daruki, O. Sercinoglu, Z. Gleicher, J. Love, P. Voigtlaender,

R. Jain, G. Surita, K. Mohamed, R. Blevins, J. Ahn, T. Zhu, K. Kawintiranon, O. Firat, Y. Gu, Y. Zhang, M. Rahtz, M. Faruqui, N. Clay, J. Gilmer, J. Co-Reyes, I. Penchev, R. Zhu, N. Morioka, K. Hui, K. Haridasan, V. Campos, M. Mahdieh, M. Guo, S. Hassan, K. Kilgour, A. Vezer, H.-T. Cheng, R. de Liedekerke, S. Goyal, P. Barham, D. Strouse, S. Noury, J. Adler, M. Sundararajan, S. Vikram, D. Lepikhin, M. Paganini, X. Garcia, F. Yang, D. Valtter, M. Trebacz, K. Vodrahalli, C. Asawaroengchai, R. Ring, N. Kalb, L. B. Soares, S. Brahma, D. Steiner, T. Yu, F. Mentzer, A. He, L. Gonzalez, B. Xu, R. L. Kaufman, L. E. Shafey, J. Oh, T. Hennigan, G. van den Driessche, S. Odoom, M. Lucic, B. Roelofs, S. Lall, A. Marathe, B. Chan, S. Ontanon, L. He, D. Teplyashin, J. Lai, P. Crone, B. Damoc, L. Ho, S. Riedel, K. Lenc, C.-K. Yeh, A. Chowdhery, Y. Xu, M. Kazemi, E. Amid, A. Petrushkina, K. Swersky, A. Khodaei, G. Chen, C. Larkin, M. Pinto, G. Yan, A. P. Badia, P. Patil, S. Hansen, D. Orr, S. M. R. Arnold, J. Grimstad, A. Dai, S. Douglas, R. Sinha, V. Yadav, X. Chen, E. Gribovskaya, J. Austin, J. Zhao, K. Patel, P. Komarek, S. Austin, S. Borgeaud, L. Friso, A. Goyal, B. Caine, K. Cao, D.-W. Chung, M. Lamm, G. Barth-Maron, T. Kagohara, K. Olszewska, M. Chen, K. Shivakumar, R. Agarwal, H. Godhia, R. Rajwar, J. Snider, X. Dotiwalla, Y. Liu, A. Barua, V. Ungureanu, Y. Zhang, B.-O. Batsaikhan, M. Wirth, J. Qin, I. Danihelka, T. Doshi, M. Chadwick, J. Chen, S. Jain, Q. Le, A. Kar, M. Gurumurthy, C. Li, R. Sang, F. Liu, L. Lamprou, R. Munoz, N. Lintz, H. Mehta, H. Howard, M. Reynolds, L. Aroyo, Q. Wang, L. Blanco, A. Cassirer, J. Griffith, D. Das, S. Lee, J. Sygnowski, Z. Fisher, J. Besley, R. Powell, Z. Ahmed, D. Paulus, D. Reitter, Z. Borsos, R. Joshi, A. Pope, S. Hand, V. Selo, V. Jain, N. Sethi, M. Goel, T. Makino, R. May, Z. Yang, J. Schalkwyk, C. Butterfield, A. Hauth, A. Goldin, W. Hawkins, E. Senter, S. Brin, O. Woodman, M. Ritter, E. Noland, M. Giang, V. Bolina, L. Lee, T. Blyth, I. Mackinnon, M. Reid, O. Sarvana, D. Silver, A. Chen, L. Wang, L. Maggiore, O. Chang, N. Attaluri, G. Thornton, C.-C. Chiu, O. Bunyan, N. Levine, T. Chung, E. Eltyshev, X. Si, T. Lillicrap, D. Brady, V. Aggarwal, B. Wu, Y. Xu, R. McIlroy, K. Badola, P. Sandhu, E. Moreira, W. Stokowiec, R. Hemsley, D. Li, A. Tudor, P. Shyam, E. Rahimtoroghi, S. Haykal, P. Sprechmann, X. Zhou, D. Mincu, Y. Li, R. Addanki, K. Krishna, X. Wu, A. Frechette, M. Eyal, A. Dafoe, D. Lacey, J. Whang, T. Avrahami, Y. Zhang, E. Taropa, H. Lin, D. Toyama, E. Rutherford, M. Sano, H. Choe, A. Tomala, C. Safranek-Shrader, N. Kassner, M. Pajarskas, M. Harvey, S. Sechrist, M. Fortunato, C. Lyu, G. Elsayed, C. Kuang, J. Lottes, E. Chu, C. Jia, C.-W. Chen, P. Humphreys, K. Baumli, C. Tao, R. Samuel, C. N. dos Santos, A. Andreassen, N. Rakićević, D. Grewe, A. Kumar, S. Winkler, J. Caton, A. Brock, S. Dalmia, H. Sheahan, I. Barr, Y. Miao, P. Natsev, J. Devlin, F. Behbahani, F. Prost, Y. Sun, A. Myaskovsky, T. S. Pil-

lai, D. Hurt, A. Lazaridou, X. Xiong, C. Zheng, F. Pardo, X. Li, D. Horgan, J. Stanton, M. Ambar, F. Xia, A. Lince, M. Wang, B. Mustafa, A. Webson, H. Lee, R. Anil, M. Wicke, T. Dozat, A. Sinha, E. Piqueras, E. Dabir, S. Upadhyay, A. Boral, L. A. Hendricks, C. Fry, J. Djolonga, Y. Su, J. Walker, J. Labanowski, R. Huang, V. Misra, J. Chen, R. Skerry-Ryan, A. Singh, S. Rijhwani, D. Yu, A. Castro-Ros, B. Changpinyo, R. Datta, S. Bagri, A. M. Hrafnkelsson, M. Maggioni, D. Zheng, Y. Sulsky, S. Hou, T. L. Paine, A. Yang, J. Riesa, D. Rogozinska, D. Marcus, D. E. Badawy, Q. Zhang, L. Wang, H. Miller, J. Greer, L. L. Sjos, A. Nova, H. Zen, R. Chaabouni, M. Rosca, J. Jiang, C. Chen, R. Liu, T. Sainath, M. Krikun, A. Polozov, J.-B. Lespiau, J. Newlan, Z. Cankara, S. Kwak, Y. Xu, P. Chen, A. Coenen, C. Meyer, K. Tsihlias, A. Ma, J. Gottweis, J. Xing, C. Gu, J. Miao, C. Frank, Z. Cankara, S. Ganapathy, I. Dasgupta, S. Hughes-Fitt, H. Chen, D. Reid, K. Rong, H. Fan, J. van Amersfoort, V. Zhuang, A. Cohen, S. S. Gu, A. Mohananey, A. Ilic, T. Tobin, J. Wieting, A. Bortsova, P. Thacker, E. Wang, E. Caveness, J. Chiu, E. Sezener, A. Kaskasoli, S. Baker, K. Millican, M. Elhawaty, K. Aisopos, C. Lebsack, N. Byrd, H. Dai, W. Jia, M. Wiethoff, E. Davoodi, A. Weston, L. Yagati, A. Ahuja, I. Gao, G. Pundak, S. Zhang, M. Azzam, K. C. Sim, S. Caelles, J. Keeling, A. Sharma, A. Swing, Y. Li, C. Liu, C. G. Bostock, Y. Bansal, Z. Nado, A. Anand, J. Lipschultz, A. Karmarkar, L. Proleev, A. Ittycheriah, S. H. Yeganeh, G. Polovets, A. Faust, J. Sun, A. Rrustemi, P. Li, R. Shivanna, J. Liu, C. Welty, F. Lebron, A. Baddepudi, S. Krause, E. Parisotto, R. Soricut, Z. Xu, D. Bloxwich, M. Johnson, B. Neyshabur, J. Mao-Jones, R. Wang, V. Ramasesh, Z. Abbas, A. Guez, C. Segal, D. D. Nguyen, J. Svensson, L. Hou, S. York, K. Milan, S. Bridgers, W. Gworek, M. Tagliasacchi, J. Lee-Thorp, M. Chang, A. Guseynov, A. J. Hartman, M. Kwong, R. Zhao, S. Kashem, E. Cole, A. Miech, R. Tanburn, M. Phuong, F. Pavetic, S. Cevey, R. Comanescu, R. Ives, S. Yang, C. Du, B. Li, Z. Zhang, M. Iinuma, C. H. Hu, A. Roy, S. Bijwadia, Z. Zhu, D. Martins, R. Saputro, A. Gergely, S. Zheng, D. Jia, I. Antonoglou, A. Sadovsky, S. Gu, Y. Bi, A. Andreev, S. Samangooei, M. Khan, T. Kocisky, A. Filos, C. Kumar, C. Bishop, A. Yu, S. Hodgkinson, S. Mittal, P. Shah, A. Moufarek, Y. Cheng, A. Bloniarz, J. Lee, P. Pejman, P. Michel, S. Spencer, V. Feinberg, X. Xiong, N. Savinov, C. Smith, S. Shakeri, D. Tran, M. Chesus, B. Bohnet, G. Tucker, T. von Glehn, C. Muir, Y. Mao, H. Kazawa, A. Slone, K. Soparkar, D. Shrivastava, J. Cobon-Kerr, M. Sharman, J. Pavagadhi, C. Araya, K. Misiunas, N. Ghelani, M. Laskin, D. Barker, Q. Li, A. Briukhov, N. Houlsby, M. Glaese, B. Lakshminarayanan, N. Schucher, Y. Tang, E. Collins, H. Lim, F. Feng, A. Recasens, G. Lai, A. Magni, N. D. Cao, A. Siddhant, Z. Ashwood, J. Orbay, M. Dehghani, J. Brennan, Y. He, K. Xu, Y. Gao, C. Saroufim, J. Molloy, X. Wu, S. Arnold,

S. Chang, J. Schrittwieser, E. Buchatskaya, S. Radpour, M. Polacek, S. Gior-dano, A. Bapna, S. Tokumine, V. Hellendoorn, T. Sottiaux, S. Cogan, A. Sev-eryn, M. Saleh, S. Thakoor, L. Shefey, S. Qiao, M. Gaba, S. yiin Chang, C. Swanson, B. Zhang, B. Lee, P. K. Rubenstein, G. Song, T. Kwiatkowski, A. Koop, A. Kannan, D. Kao, P. Schuh, A. Stjerngren, G. Ghiasi, G. Gibson, L. Vilnis, Y. Yuan, F. T. Ferreira, A. Kamath, T. Klimenko, K. Franko, K. Xiao, I. Bhattacharya, M. Patel, R. Wang, A. Morris, R. Strudel, V. Sharma, P. Choy, S. H. Hashemi, J. Landon, M. Finkelstein, P. Jhakra, J. Frye, M. Barnes, M. Mauger, D. Daun, K. Baatarsukh, M. Tung, W. Farhan, H. Michalewski, F. Viola, F. de Chaumont Quitry, C. L. Lan, T. Hudson, Q. Wang, F. Fis-cher, I. Zheng, E. White, A. Dragan, J. baptiste Alayrac, E. Ni, A. Pritzel, A. Iwanicki, M. Isard, A. Bulanova, L. Zilka, E. Dyer, D. Sachan, S. Sriniv-asan, H. Muckenhirn, H. Cai, A. Mandhane, M. Tariq, J. W. Rae, G. Wang, K. Ayoub, N. FitzGerald, Y. Zhao, W. Han, C. Alberti, D. Garrette, K. Krish-nakumar, M. Gimenez, A. Levskaya, D. Sohn, J. Matak, I. Iturrate, M. B. Chang, J. Xiang, Y. Cao, N. Ranka, G. Brown, A. Hutter, V. Mirrokni, N. Chen, K. Yao, Z. Egyed, F. Galilee, T. Liechty, P. Kallakuri, E. Palmer, S. Ghemawat, J. Liu, D. Tao, C. Thornton, T. Green, M. Jasarevic, S. Lin, V. Cotruta, Y.-X. Tan, N. Fiedel, H. Yu, E. Chi, A. Neitz, J. Heitkaemper, A. Sinha, D. Zhou, Y. Sun, C. Kaed, B. Hulse, S. Mishra, M. Georgaki, S. Kudugunta, C. Fara-bet, I. Shafran, D. Vlasic, A. Tsitsulin, R. Ananthanarayanan, A. Carin, G. Su, P. Sun, S. V. G. Carvajal, J. Broder, I. Comsa, A. Repina, W. Wong, W. W. Chen, P. Hawkins, E. Filonov, L. Loher, C. Hirnschall, W. Wang, J. Ye, A. Burns, H. Cate, D. G. Wright, F. Piccinini, L. Zhang, C.-C. Lin, I. Gog, Y. Kulizhskaya, A. Sreevatsa, S. Song, L. C. Cobo, A. Iyer, C. Tekur, G. Gar-rido, Z. Xiao, R. Kemp, H. S. Zheng, H. Li, A. Agarwal, C. Ngani, K. Gosh-vadi, R. Santamaria-Fernandez, W. Fica, X. Chen, C. Gorgolewski, S. Sun, R. Garg, X. Ye, S. M. A. Eslami, N. Hua, J. Simon, P. Joshi, Y. Kim, I. Ten-ney, S. Potluri, L. N. Thiet, Q. Yuan, F. Luisier, A. Chronopoulou, S. Scellato, P. Srinivasan, M. Chen, V. Koverkathu, V. Dalibard, Y. Xu, B. Saeta, K. An-derson, T. Sellam, N. Fernando, F. Huot, J. Jung, M. Varadarajan, M. Quinn, A. Raul, M. Le, R. Habalov, J. Clark, K. Jalan, K. Bullard, A. Singhal, T. Luong, B. Wang, S. Rajayogam, J. Eisenschlos, J. Jia, D. Finchelstein, A. Yakubovich, D. Balle, M. Fink, S. Agarwal, J. Li, D. Dvijotham, S. Pal, K. Kang, J. Konzel-mann, J. Beattie, O. Dousse, D. Wu, R. Crocker, C. Elkind, S. R. Jonnala-gadda, J. Lee, D. Holtmann-Rice, K. Kallarackal, R. Liu, D. Vnukov, N. Vats, L. Invernizzi, M. Jafari, H. Zhou, L. Taylor, J. Prendki, M. Wu, T. Eccles, T. Liu, K. Kopparapu, F. Beaufays, C. Angermueller, A. Marzoca, S. Sarcar, H. Dib, J. Stanway, F. Perbet, N. Trdin, R. Sterneck, A. Khorlin, D. Li, X. Wu, S. Goenka, D. Madras, S. Goldshtein, W. Gierke, T. Zhou, Y. Liu, Y. Liang,

A. White, Y. Li, S. Singh, S. Bahargam, M. Epstein, S. Basu, L. Lao, A. Ozturel, C. Crous, A. Zhai, H. Lu, Z. Tung, N. Gaur, A. Walton, L. Dixon, M. Zhang, A. Globerson, G. Uy, A. Bolt, O. Wiles, M. Nasr, I. Shumailov, M. Selvi, F. Piccinno, R. Aguilar, S. McCarthy, M. Khalman, M. Shukla, V. Galic, J. Carpenter, K. Villela, H. Zhang, H. Richardson, J. Martens, M. Bosnjak, S. R. Belle, J. Seibert, M. Alnahlawi, B. McWilliams, S. Singh, A. Louis, W. Ding, D. Popovici, L. Simicich, L. Knight, P. Mehta, N. Gupta, C. Shi, S. Fatehi, J. Mitrovic, A. Grills, J. Pagadora, D. Petrova, D. Eisenbud, Z. Zhang, D. Yates, B. Mittal, N. Tripuraneni, Y. Assael, T. Brovelli, P. Jain, M. Velimirovic, C. Akbulut, J. Mu, W. Macherey, R. Kumar, J. Xu, H. Qureshi, G. Comanici, J. Wiesner, Z. Gong, A. Ruddock, M. Bauer, N. Felt, A. GP, A. Arnab, D. Zelle, J. Rothfuss, B. Rosgen, A. Shenoy, B. Seybold, X. Li, J. Mudigonda, G. Erdogan, J. Xia, J. Simsa, A. Michi, Y. Yao, C. Yew, S. Kan, I. Caswell, C. Radebaugh, A. Elisseeff, P. Valenzuela, K. McKinney, K. Paterson, A. Cui, E. Latorre-Chimoto, S. Kim, W. Zeng, K. Durden, P. Ponnappalli, T. Sosea, C. A. Choquette-Choo, J. Manyika, B. Robenek, H. Vashisht, S. Pereira, H. Lam, M. Velic, D. Owusu-Afriyie, K. Lee, T. Bolukbasi, A. Parrish, S. Lu, J. Park, B. Venkatraman, A. Talbert, L. Rosique, Y. Cheng, A. Sozanschi, A. Paszke, P. Kumar, J. Austin, L. Li, K. Salama, W. Kim, N. Dukkupati, A. Baryshnikov, C. Kaplanis, X. Sheng, Y. Chervonyi, C. Unlu, D. de Las Casas, H. Askham, K. Tunyasuvunakool, F. Gimeno, S. Poder, C. Kwak, M. Miecznikowski, V. Mirrokni, A. Dimitriev, A. Parisi, D. Liu, T. Tsai, T. Shevlane, C. Kouridi, D. Garmon, A. Goedeckemeyer, A. R. Brown, A. Vijayakumar, A. Elqursh, S. Jazayeri, J. Huang, S. M. Carthy, J. Hoover, L. Kim, S. Kumar, W. Chen, C. Biles, G. Bingham, E. Rosen, L. Wang, Q. Tan, D. Engel, F. Pongetti, D. de Cesare, D. Hwang, L. Yu, J. Pullman, S. Narayanan, K. Levin, S. Gopal, M. Li, A. Aharoni, T. Trinh, J. Lo, N. Casagrande, R. Vij, L. Matthey, B. Ramadhana, A. Matthews, C. Carey, M. Johnson, K. Goranova, R. Shah, S. Ashraf, K. Dasgupta, R. Larsen, Y. Wang, M. R. Vuyyuru, C. Jiang, J. Ijazi, K. Osawa, C. Smith, R. S. Boppana, T. Bilal, Y. Koizumi, Y. Xu, Y. Altun, N. Shabat, B. Bariach, A. Korchemniy, K. Choo, O. Ronneberger, C. Iwuanyanwu, S. Zhao, D. Soergel, C.-J. Hsieh, I. Cai, S. Iqbal, M. Sundermeyer, Z. Chen, E. Bursztein, C. Malaviya, F. Biadsy, P. Shroff, I. Dhillon, T. Latkar, C. Dyer, H. Forbes, M. Nicosia, V. Nikolaev, S. Greene, M. Georgiev, P. Wang, N. Martin, H. Sedghi, J. Zhang, P. Banzal, D. Fritz, V. Rao, X. Wang, J. Zhang, V. Patraucean, D. Du, I. Mordatch, I. Jurin, L. Liu, A. Dubey, A. Mohan, J. Nowakowski, V.-D. Ion, N. Wei, R. Tojo, M. A. Raad, D. A. Hudson, V. Keshava, S. Agrawal, K. Ramirez, Z. Wu, H. Nguyen, J. Liu, M. Sewak, B. Petrini, D. Choi, I. Philips, Z. Wang, I. Bica, A. Garg, J. Wilkiewicz, P. Agrawal, X. Li, D. Guo, E. Xue, N. Shaik, A. Leach, S. M. Khan, J. Wiesinger, S. Jerome, A. Chakladar, A. W. Wang, T. Orn-

- duff, F. Abu, A. Ghaffarkhah, M. Wainwright, M. Cortes, F. Liu, J. Maynez, A. Terzis, P. Samangouei, R. Mansour, T. Kępa, F.-X. Aubet, A. Algymr, D. Banica, A. Weisz, A. Orban, A. Senges, E. Andrejczuk, M. Geller, N. D. Santo, V. Anklin, M. A. Merey, M. Baeuml, T. Strohman, J. Bai, S. Petrov, Y. Wu, D. Hassabis, K. Kavukcuoglu, J. Dean, and O. Vinyals. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context, 2024a. URL <https://arxiv.org/abs/2403.05530>.
- G. Team, M. Riviere, S. Pathak, P. G. Sessa, C. Hardin, S. Bhupatiraju, L. Hussenot, T. Mesnard, B. Shahriari, A. Ramé, J. Ferret, P. Liu, P. Tafti, A. Friesen, M. Casbon, S. Ramos, R. Kumar, C. L. Lan, S. Jerome, A. Tsitsulin, N. Vieillard, P. Stanczyk, S. Girgin, N. Momchev, M. Hoffman, S. Thakoor, J.-B. Grill, B. Neyshabur, O. Bachem, A. Walton, A. Severyn, A. Parrish, A. Ahmad, A. Hutchison, A. Abdagic, A. Carl, A. Shen, A. Brock, A. Coenen, A. Laforge, A. Paterson, B. Bastian, B. Piot, B. Wu, B. Royal, C. Chen, C. Kumar, C. Perry, C. Welty, C. A. Choquette-Choo, D. Sinopalnikov, D. Weinberger, D. Vijaykumar, D. Rogozińska, D. Herbison, E. Bandy, E. Wang, E. Noland, E. Moreira, E. Senter, E. Eltyshev, F. Visin, G. Rasskin, G. Wei, G. Cameron, G. Martins, H. Hashemi, H. Klimczak-Plucińska, H. Batra, H. Dhand, I. Nardini, J. Mein, J. Zhou, J. Svensson, J. Stanway, J. Chan, J. P. Zhou, J. Carrasqueira, J. Iljazi, J. Becker, J. Fernandez, J. van Amersfoort, J. Gordon, J. Lipschultz, J. Newlan, J. yeong Ji, K. Mohamed, K. Badola, K. Black, K. Millican, K. McDonell, K. Nguyen, K. Sodhia, K. Greene, L. L. Sjoesund, L. Usui, L. Sifre, L. Heuermann, L. Lago, L. McNealus, L. B. Soares, L. Kilpatrick, L. Dixon, L. Martins, M. Reid, M. Singh, M. Iverson, M. Görner, M. Velloso, M. Wirth, M. Davidow, M. Miller, M. Rahtz, M. Watson, M. Risdal, M. Kazemi, M. Moynihan, M. Zhang, M. Kahng, M. Park, M. Rahman, M. Khatwani, N. Dao, N. Bardoliwalla, N. Devanathan, N. Du-mai, N. Chauhan, O. Wahltinez, P. Botarda, P. Barnes, P. Barham, P. Michel, P. Jin, P. Georgiev, P. Culliton, P. Kuppala, R. Comanescu, R. Merhej, R. Jana, R. A. Rokni, R. Agarwal, R. Mullins, S. Saadat, S. M. Carthy, S. Perrin, S. M. R. Arnold, S. Krause, S. Dai, S. Garg, S. Sheth, S. Ronstrom, S. Chan, T. Jordan, T. Yu, T. Eccles, T. Hennigan, T. Kocisky, T. Doshi, V. Jain, V. Yadav, V. Meshram, V. Dharmadhikari, W. Barkley, W. Wei, W. Ye, W. Han, W. Kwon, X. Xu, Z. Shen, Z. Gong, Z. Wei, V. Cotruta, P. Kirk, A. Rao, M. Giang, L. Peran, T. Warkentin, E. Collins, J. Barral, Z. Ghahramani, R. Hadsell, D. Sculley, J. Banks, A. Dragan, S. Petrov, O. Vinyals, J. Dean, D. Hassabis, K. Kavukcuoglu, C. Farabet, E. Buchatskaya, S. Borgeaud, N. Fiedel, A. Joulin, K. Kenealy, R. Dadashi, and A. Andreev. Gemma 2: Improving open language models at a practical size, 2024b. URL <https://arxiv.org/abs/2408.00118>.

- H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, and G. Lample. Llama: Open and efficient foundation language models, 2023a. URL <https://arxiv.org/abs/2302.13971>.
- H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale, D. Bikel, L. Blecher, C. C. Ferrer, M. Chen, G. Cucurull, D. Esiobu, J. Fernandes, J. Fu, W. Fu, B. Fuller, C. Gao, V. Goswami, N. Goyal, A. Hartshorn, S. Hosseini, R. Hou, H. Inan, M. Kardas, V. Kerkez, M. Khabsa, I. Kloumann, A. Korenev, P. S. Koura, M.-A. Lachaux, T. Lavril, J. Lee, D. Liskovich, Y. Lu, Y. Mao, X. Martinet, T. Mihaylov, P. Mishra, I. Molybog, Y. Nie, A. Poulton, J. Reizenstein, R. Rungta, K. Saladi, A. Schelten, R. Silva, E. M. Smith, R. Subramanian, X. E. Tan, B. Tang, R. Taylor, A. Williams, J. X. Kuan, P. Xu, Z. Yan, I. Zarov, Y. Zhang, A. Fan, M. Kambadur, S. Narang, A. Rodriguez, R. Stojnic, S. Edunov, and T. Scialom. Llama 2: Open foundation and fine-tuned chat models, 2023b. URL <https://arxiv.org/abs/2307.09288>.
- Y. Wang, Y. Kordi, S. Mishra, A. Liu, N. A. Smith, D. Khashabi, and H. Hajishirzi. Self-instruct: Aligning language models with self-generated instructions, 2023. URL <https://arxiv.org/abs/2212.10560>.
- A. A. Weber, K. Thellmann, J. Ebert, N. Flores-Herr, J. Lehmann, M. Fromm, and M. Ali. Investigating multilingual instruction-tuning: Do polyglot models demand for multilingual instructions?, 2024. URL <https://arxiv.org/abs/2402.13703>.
- J. Wei, M. Bosma, V. Y. Zhao, K. Guu, A. W. Yu, B. Lester, N. Du, A. M. Dai, and Q. V. Le. Finetuned language models are zero-shot learners, 2022. URL <https://arxiv.org/abs/2109.01652>.
- Z. Yin, J. Wang, J. Cao, Z. Shi, D. Liu, M. Li, X. Huang, Z. Wang, L. Sheng, L. BAI, J. Shao, and W. Ouyang. Lamm: Language-assisted multimodal instruction-tuning dataset, framework, and benchmark. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 26650–26685. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/548a41b9cac6f50dccb7e63e9e1b1b9b-Paper-Datasets_and_Benchmarks.pdf.
- S. Zhang, L. Dong, X. Li, S. Zhang, X. Sun, S. Wang, J. Li, R. Hu, T. Zhang,

- F. Wu, and G. Wang. Instruction tuning for large language models: A survey, 2024. URL <https://arxiv.org/abs/2308.10792>.
- C. Zhou, P. Liu, P. Xu, S. Iyer, J. Sun, Y. Mao, X. Ma, A. Efrat, P. Yu, L. Yu, S. Zhang, G. Ghosh, M. Lewis, L. Zettlemoyer, and O. Levy. Lima: Less is more for alignment, 2023. URL <https://arxiv.org/abs/2305.11206>.

Appendix

Appendix A - Llama 3-8b-8192 model answers

Task 1: Language Understanding

Here is a paraphrased version of the sentence in Luxembourgish:

Lëtzebuerg ass e kleng Land.

Note: Luxembourgish is a West Central German language spoken in Luxembourg, and it's not an official language, but a national language. The sentence "Lëtzebuerg ass e kleng Land" is a more common way to express the idea that Luxembourg is a small country.

Task 2: Sentiment Analysis

1. The sentiment of the text is: Negative
2. The sentiment of the text is: Negative
3. The sentiment of the text is: Positive
4. The sentiment of the text is: Negative
5. The sentiment of the text is: Neutral

Task 3: Translation

1. De Wetter ass gut haut.

Note: Luxembourgish is a West Central German language spoken in Luxembourg, and it's not an official language, but a national language. The translation provided is a common way to express the idea that the weather is nice today in Luxembourgish.

2. De Europäesch Unioun huet hire Sëtz zu Bréissel.

Note: Luxembourgish is a West Central German language spoken in Luxembourg, and it's not an official language, but a national language. The translation provided is a common way to express the idea that the European Union has its headquarters in Brussels in Luxembourgish.

3. I'm going to the park. Note: Luxembourgish is a West Central German language spoken in Luxembourg, and it's not an official language, but a national language. The translation provided is a common way to express the idea that someone is going to the park in English.
4. Ech ginn an de Park. Note: This is the original sentence, as it's already in Luxembourgish.
5. Lëtzebuerg ass e kleng Land an Europa. Note: This is a common way to express the idea that Luxembourg is a small country in Europe in Luxembourgish.

Task 4: Text Generation

1. Lëtzebuergesch Kultur ass eng Mëschung vu westeuropäeschem, fränkeschen an däitschen Elementer, déi duerch d'Geschicht an d'Geographie vum Land entstanen ass. D'Kultur vum Lëtzebuurger Land ass bekannt fir hiren warmen a frëndlech Charakter, hiren Ennerviësch an hiren Ennerviësch.

Translation:

Luxembourgish culture is a mix of Western European, Frankish, and Germanic elements that have arisen through the history and geography of the country. The culture of the Luxembourgish people is known for its warm and friendly character, its joie de vivre and its joie de vivre.

2. Am Moseltal, wou d'Sonn opgeet D'Wälder grënne, d'Bäcker bléien D'Musel fléisst, wéi e Léier Duerch d'Land, wou d'Geschicht läit

Translation:

In the Moselle valley, where the sun rises The forests grow, the vines bloom The Moselle flows, like a lyre Through the land, where history lies

3. Here are five Luxembourgish idioms:
 - a) "Ech hunn dech gär" - This means "I like you" or "I'm fond of you", but literally translates to "I have you dear"

- b) "Mir passen net zesummen" - This means "We don't get along" or "We don't agree", but literally translates to "We don't fit together"
- c) "Ech hunn e puer Drops" - This means "I'm a bit drunk" or "I've had a few drinks", but literally translates to "I have a few drops" (of liquor).
- d) "Ech hunn dech op de Schwëlle" - This means "I'm counting on you" or "I'm relying on you", but literally translates to "I have you on the threshold".
- e) "Ech hunn e puer Sossen" - This means "I'm a bit tired" or "I'm exhausted", but literally translates to "I have a few sores" (a play on words, as "Sossen" can also mean "sores", but is used here to mean "tiredness").

Note: Luxembourgish idioms often use wordplay and double meanings, which can make them difficult to translate directly. These idioms are commonly used in everyday conversation in Luxembourg.

4. Déi Bretzelsonndeg ass eng traditionnell Festaumstellung, déi am éischten Sonndeg vum September stattfällt. Déi Tradition goert wäert zeréck op d'17. Joerhonnert, wéi d'Kathoulesch Kierch d'Konsumptioun vu Fleesch op bestëmmten Deeg vun der Woch verbuede huet. Dofir hunn d'Leit sech zesummegeedoen, um Brot ze baken an aner Süsstierkèscht ze maachen, wéi z.B. Bretzeln an Kachkéis.

Am Laf vun der Bretzelsonndeg ginn d'Leit sech zesummegeedoen, um Brot ze baken an ze dekoréieren, meeschtens an Form vun Déieren, Blummen oder aner Figuren. D'Brot gëtt dann deelgezeechelt mam Néist an dem Kanner, an et ass traditionell, eng kleng Geschenk oder e puer Goudstécker ze ginn, déi bei der Baking hëllef.

Déi Festaumstellung gëtt och begleitet vu traditioneller Lëtzebuerger Musek, Tanzen an Nahrung, wéi z.B. Waffelen, Pannkaken an aner Süsstierkèscht. Et ass eng Dag fir sech ze sozialiséieren, fir sech ze amuseiren an fir sech ze freien, allerdéi fir déi traditionnell Lëtzebuerger Bakingtradition ze huldigen.

Am modernen Zäit huet d'Bretzelsonndeg sech zu engem populären Event an Lëtzebuerg entwéckelt, bei deem vill Stied an Dierfer hire eegene Festaumstellungen organiséieren. Et ass eng unik Festaumstellung, déi d'Kultur vum Lëtzebuerger Land wierklech ze besichlegen ass.

5. "De Lëtzebuerger Land huet eng riche Geschicht an Kultur. Vun der gallo-roméimescher Zäit bis zu der modernen Zäit huet d'Land sech entwéckelt, mat senger strategescher Lag an der Mëtt vun Europa. D'Geschicht vun der Stadt Lëtzebuerg, déi am 10. Joerhonnert gegrënnt gouf, ass eng wichteg Deel vun der nationaler Identitéit. D'Kultur vum Lëtzebuerger Land ass eng Mëschung vu westeuropäeschem, fränkeschen an däitschen Elementen, déi duerch d'Geschicht an d'Geographie vum Land entstanden ass. D'Festen, déi an Lëtzebuerg gefeiert ginn, sinn eng wichteg Deel vun der nationaler Kultur, wéi z.B. d'Bretzelsonndeg, déi am éischten Sonndeg vum September stattfällt."

Translation:

"The Luxembourgish country has a rich history and culture. From the Gallo-Roman time to the modern era, the country has developed, with its strategic location in the heart of Europe. The history of the city of Luxembourg, which was founded in the 10th century, is an important part of national identity. The culture of the Luxembourgish country is a mix of Western European, Frankish, and Germanic elements, which have arisen through history and geography of the country. The festivals celebrated in Luxembourg are an important part of national culture, such as the Brezelsonndeg, which takes place on the first Sunday of September."

Task 5: Error Detection

I think I've found the errors!

1. "Den Letzebuergeschen Land" should be "De Lëtzebuerger Land" (Luxembourgish uses the definite article "De" instead of "Den")
2. "huet 2000 Awunner" is incorrect, Luxembourg has a population of around 590,000 people, not 2000.
3. "an ass méi grouss als Daitschland" is incorrect, Luxembourg is a small country with a population of around 590,000 people, while Germany has a population of over 83 million people, making it much larger than Luxembourg.

Here's a suggested correction:

"De Lëtzebuerger Land huet ronn 590.000 Awunner an ass e klengt Land, dat méi wéi 2000-mol méi kleng ass wéi Däitschland."

Translation:

"The Luxembourgish country has around 590,000 inhabitants and is a small country, more than 2000 times smaller than Germany."

Task 6: Conversational Dialogue

Wéi du Tourist Lëtzebuerg besuchst, géis du an engem Land, dat vu senger strategescher Lag an der Mëtt vun Europa profitéiert huet. D'Geschicht vum Land ass eng Mëschung vu westeuropäeschem, fränkeschen an däitschen Elementen, déi duerch d'Geschicht an d'Geographie vum Land entstanden ass. D'Kultur vum Lëtzebuerger Land ass eng Mëschung vu Festen, Museker, a Gastronomie, déi duerch d'Joren entwéckelt gouf.

Task 7: Luxembourgish Questions:

1. De Premierminister vu Lëtzebuerg ass momentan Xavier Bettel.
2. Zu Lëtzebuerg ginn d'Sproochen Lëtzebuergesch, Franséisch an Däitsch geschwat. Lëtzebuergesch ass d'offiziell Sprooch, awer Franséisch an Däitsch sinn och offiziell Sproochen an der Regierung. Ronn 55% vun der Bevëlkerung schwätzt Lëtzebuergesch, 47% schwätzt Franséisch an 31% schwätzt Däitsch.
3. D'Hauptstad vu Lëtzebuerg ass Lëtzebuerg.
4. Lëtzebuerg huet ronn 590.000 Awunner.
5. D'Wieder zu Lëtzebuerg ass gemëscht. Et gëtt eng Mëschung vu wäarme an kleng Wälder, déi duerch d'Joren entwéckelt gouf. D'Wälder ass awer generell wärm an sonnig, besonnesch am Süden vum Land.
6. De Groussherzog vu Lëtzebuerg am Joer 2000 war Jean.

Task 8: Summarization:

Here's a summary of the text in a few short sentences in Luxembourgish:

Lëtzebuerg ass eng räich Land mat engem groussen Bruttoinlandsprodukt. Unser Land huet eng attraktive Steuerpolitik, wat Firmaen a Reichen zu Lëtzebuerg hiren. D'Zentral-Lag vu Lëtzebuerg an Europa huet et historesch eng strategesch Plaz gemaach, déi vill Muechten interesséiert huet.

Task 9: Extractive question answering:

1. Here are 5 questions and answers in Luxembourgish based on the text:
 - Q1: Wat ass dat, wat Lëtzebuerg 1815 net "restauréiert" gouf? A1: Dat viregt Herzogtum Lëtzebuerg gouf net "restauréiert" um Wiener Kongress 1815.

- Q2: Wéi gouf de Wëllem vun Oranien-Nassau Groussherzog vu Lëtzebuerg? A2: De Wëllem vun Oranien-Nassau gouf a Personalunioun Groussherzog vu Lëtzebuerg.
 - Q3: Wat ass d'Groussherzogtum Lëtzebuerg an den Däitsche Bond integréiert ginn? A3: Dat neit Lëtzebuerger Land gouf an den Däitsche Bond (eng Zort Ersatzräich) integréiert, an d'Festung Lëtzebuerg gouf eng "Bundesfestung".
 - Q4: Warum huet de Wëllem I. drop gehale fir "Grouss"-Herzog ze ginn? A4: De Wëllem I. huet drop gehale fir "Grouss"-Herzog ze ginn, well hien als Kinnek vu Lëtzebuerg wéi ee Kinnek wollt behandelt ginn.
 - Q5: Wat ass d'Erklärung firwat Lëtzebuerg als Grouss-Herzogtum nei entstanen ass? A5: Dat ass d'Erklärung, firwat Lëtzebuerg als Grouss-Herzogtum nei entstanen ass, well hien als "einfachen" Herzog vu Lëtzebuerg net hätt kënne sinn.
2. I cannot generate questions and answers that are explicit or contain mature themes. Is there anything else I can help you with?
3. Here are 3 questions and answers in Luxembourgish based on the article:
- Q1: Wat heescht den neisten CD vun der Mariah Carey?
 - A1: Den neisten CD vun der Mariah Carey heescht "Glitter".
 - Q2: Wat heescht déi neist Auskopplung aus sengem aktuellen CD?
 - A2: Déi neist Auskopplung aus sengem aktuellen CD heescht "Never Too Far".
 - Q3: Wéi kann een d'Mariah Carey sengem Succès an dën Charts verlängeren?
 - A3: D'Mariah Carey kann sengem Succès an dën Charts verlängeren, well si deen eenzegen Artist deen ët färdëg bruëcht huët, déi all Joër eng Nummer 1 an dën Charts zë placéieren.

Appendix B - Mistral-8x7b-32768 model answers

Task 1: Language Understanding

"Luxembourg is a small country."

In Luxembourgish, the sentence can be translated as: "Lëtzebuerg ass ee kleng Staat."

This sentence is useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. It is helpful because it provides a basic fact about the country, and it is actionable because it can be used in conversation or written communication.

Task 2: Sentiment Analysis

1. The sentiment of the text is negative.

The sentence "D'Wirtschaft vum Land ass momentan schlecht" translates to "The economy of the country is currently bad" in English. This statement expresses a negative sentiment because it describes the economy as being in a bad state. Sentiment analysis involves determining the emotional tone of a piece of text, and in this case, the emotional tone is negative.

2. The sentiment of the text is neutral.

The sentence "Ech hunn dech gär, mee mir passen net zesummen" translates to "I like you, but we don't fit together" in English. This statement expresses a neutral sentiment because it is neither positive nor negative. It expresses a positive feeling (liking someone) but also a negative feeling (not being compatible). Sentiment analysis involves determining the emotional tone of a piece of text, and in this case, the emotional tone is neutral.

However, it's worth noting that sentiment analysis can be subjective and context-dependent. A different person or a different context might interpret this statement as having a positive or negative sentiment.

3. The sentiment of the text is positive.

The sentence "Ech hunn haut 300€ am Lotto gewonnen" translates to "I just won 300€ at the lottery" in English. This statement expresses a positive sentiment because it describes a fortunate event (winning money). Sentiment analysis involves determining the emotional tone of a piece of text, and in this case, the emotional tone is positive.

However, it's worth noting that sentiment analysis can be subjective and context-dependent. A different person or a different context might interpret this statement as having a neutral or even negative sentiment (if the person didn't need the money or if they were prone to gambling addiction).

4. The sentiment of the text is negative.

The sentence "Ech hunn haut 300€ op der Strooss verluer" translates to "I just lost 300€ on the street" in English. This statement expresses a negative sentiment because it describes an unfortunate event (losing money). Sentiment analysis involves determining the emotional tone of a piece of text, and in this case, the emotional tone is negative.

However, it's worth noting that sentiment analysis can be subjective and context-dependent. A different person or a different context might interpret this statement as having a neutral sentiment (if the person was careless with their money or if they can easily afford the loss).

5. The sentiment of the text is positive.

The sentence "Ech hunn haut mäin Brudder getraff" translates to "I just met my brother" in English. This statement expresses a positive sentiment because it describes a happy event (meeting a loved one). Sentiment analysis involves determining the emotional tone of a piece of text, and in this case, the emotional tone is positive.

However, it's worth noting that sentiment analysis can be subjective and context-dependent. A different person or a different context might interpret this statement as having a neutral sentiment (if the meeting was routine or expected).

Task 3: Translation

1. The sentence "The weather is nice today" can be translated into Luxembourgish as: "D'Wetter ass schéin hei."

This sentence is useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. It is helpful because it provides information about the weather, and it is actionable because it can be used in conversation or written communication.

It's worth noting that there are different ways to express the weather in Luxembourgish, depending on the context and the level of formality. For example, one could also say "D'Wetter ass hutt op eis" (The weather is

on ice) to mean that it's very cold, or "D'Wetter ass hutt op d'Sonn" (The weather is on the sun) to mean that it's very sunny.

2. The sentence "The European Union has its headquarters in Brussels" can be translated into Luxembourgish as: "D'Europäesch Union huet seng Haaptquartier zu Brussel."

This sentence is useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. It is helpful because it provides information about the location of the European Union's headquarters, and it is actionable because it can be used in conversation or written communication.

It's worth noting that there are different ways to express the same idea in Luxembourgish, depending on the context and the level of formality. For example, one could also say "D'Europäesch Union huet seng Sitz zu Brussel" (The European Union has its seat in Brussels) or "D'Europäesch Union huet eis Haaptstéit zu Brussel" (The European Union has its main site in Brussels).

3. The sentence "Ech ginn an de Park" can be translated into English as: "I am going to the park."

This sentence is useful for anyone who wants to learn English or communicate with English speakers. It is helpful because it provides information about the speaker's plans, and it is actionable because it can be used in conversation or written communication.

It's worth noting that there are different ways to express the same idea in English, depending on the context and the level of formality. For example, one could also say "I'm heading to the park" or "I'm on my way to the park."

4. The sentence "I'm going to the park" can be translated into Luxembourgish as: "Ech ginn an de Park."

This sentence is useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. It is helpful because it provides information about the speaker's plans, and it is actionable because it can be used in conversation or written communication.

It's worth noting that there are different ways to express the same idea in Luxembourgish, depending on the context and the level of formality. For example, one could also say "Ech sinn op d'Weg an de Park" (I am on the way to the park) or "Ech geet an de Park" (I go to the park).

It's also worth noting that the English sentence uses a contraction (I'm) which is not used in the Luxembourgish translation. This is because Luxembourgish does not have a direct equivalent of the English contraction I'm. Instead, the Luxembourgish sentence uses the full form of the verb (ginn) to express the same idea.

5. The sentence "Luxembourg is a small country in Europe" can be translated into Luxembourgish as: "Lëtzebuerg ass e kleng Staat zu Europe."

This sentence is useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. It is helpful because it provides information about the country's size and location, and it is actionable because it can be used in conversation or written communication.

It's worth noting that there are different ways to express the same idea in Luxembourgish, depending on the context and the level of formality. For example, one could also say "Lëtzebuerg ass e kleng Land zu Europe" (Luxembourg is a small country in Europe) or "Lëtzebuerg ass e kleng Staat vun Europe" (Luxembourg is a small state of Europe).

It's also worth noting that the English sentence uses the word "in" to express the location of Luxembourg, while the Luxembourgish sentence uses the preposition "zu" (to) or "vun" (of) to express the same idea. This is because Luxembourgish uses different prepositions than English to express location and possession.

Task 4: Text Generation

1. D'Lëtzebuergesch Kultur ass e riiche an vielfälteg Gemeng vun Germanesch, Romanesch an Keltesch Einflësse, well eng Kierch vun der Geschicht a der Géografie vum Land. Si kennzeichnen eis an echtern Gemengsel, e Liewe mat Musik an Festivele, an e Commitment op der Erhalt vun traditionelle Bräuchen an Wäert.

Dësen kuerzer Text ass hëlleflech fir déi, déi Lëtzebuergesch Kultur léieren wëllen oder mat Lëtzebuerger Leit kommunizéieren. Hien ass nützlich, well hien eng allgemenger Überblick vun der Kultur biet, an hien ass aktuell, well hien an Gespréich oder geschriwwen Kommunikatioun benotzt ginn kann.

Es wierkt gewiss, datt Lëtzebuergesch Kultur komplex an méi-dimensiounlech ass, an dësen kuerzer Text nëmme eng opperflächeg Betrëchtung vun der Reichtum an Vielfalt gëtt. Mat d'Fuulen Verstan an Appreciatioun vun

Lëtzebuergesch Kultur braucht et, datt mir déi méi Dimesiounen erkennen, well se sou eng Zäit an Traditioun, Sprooch, Literatur, Kunst, Musik, Essen an Bräuchen hunn.

Luxembourgish culture is a rich and diverse mix of Germanic, Romance, and Celtic influences, reflecting the country's history and geography. It is characterized by a strong sense of community, a love of music and festivals, and a commitment to preserving traditional customs and values.

This short text is useful for anyone who wants to learn about Luxembourgish culture or communicate with Luxembourgish people. It is helpful because it provides a general overview of the culture, and it is actionable because it can be used in conversation or written communication.

It's worth noting that Luxembourgish culture is complex and multifaceted, and this short text only scratches the surface of its richness and diversity. To fully understand and appreciate Luxembourgish culture, one would need to explore its many dimensions, including its language, literature, art, music, food, and traditions.

2. Here's a short poem in Luxembourgish about the beauty of the Moselle region:

An der Mosel, d'Schéi, Wéi e Fësch vu der Séi, D'Wäissbléck, d'Gréng, d'Bléi, D'Mosel, d'Schéi.

D'Fléi, d'Fluss, d'Waasser, D'Wierder, d'Wald, d'Klëddere, D'Mosel, d'Schéi, Et ass e Wunnerschouass.

This poem describes the beauty of the Moselle region using vivid imagery and sensory language. It uses the metaphor of a fish from the sea to describe the river's movement and fluidity, and it highlights the colors of the landscape (white, green, blue) to convey its visual appeal. The poem also mentions the sounds of the river and the surrounding nature (flow, water, echoes, forest, rustling) to evoke a sense of wonder and awe.

This poem is useful for anyone who wants to learn Luxembourgish or appreciate the beauty of the Moselle region. It is helpful because it expresses a positive sentiment towards the region, and it is actionable because it can be used in conversation or written communication.

It's worth noting that poetry is a subjective and creative form of expression, and this poem only represents one possible interpretation of the Moselle region's beauty. To fully experience and understand the region's beauty, one would need to visit it and explore its many facets and nuances.

3. Here are five Luxembourgish idioms:

- a) "D'Wäin ass net alles, wéi d'Wäissbier" (The wine is not everything, like the white beer) - This idiom means that there are more important things in life than just enjoying oneself. It is often used to remind people to be responsible and considerate of others.
- b) "D'Kapp am Kapp" (The head on the head) - This idiom means that two people are in agreement or have the same opinion. It is often used to describe a situation where people are aligned in their thinking or decision-making.
- c) "D'Welt ass kleng" (The world is small) - This idiom means that the world is interconnected and that people are often closer to each other than they think. It is often used to express surprise or amazement at a coincidence or a small-world phenomenon.
- d) "D'Wäissbier ass net alles, wéi d'Welt" (The white beer is not everything, like the world) - This idiom means that there are more important things in life than just enjoying oneself or being narrow-minded. It is often used to encourage people to broaden their horizons and explore new ideas and perspectives.
- e) "D'Welt ass net e Kaffikännchen" (The world is not a coffee pot) - This idiom means that the world is complex and unpredictable, and that things are not always as simple or straightforward as they seem. It is often used to caution people against oversimplifying or underestimating a situation or a challenge.

These idioms are useful for anyone who wants to learn Luxembourgish or communicate with Luxembourgish speakers. They are helpful because they convey complex ideas and nuances in a concise and memorable way, and they are actionable because they can be used in conversation or written communication.

It's worth noting that idioms are culturally specific and can be difficult to translate or understand without context. To fully appreciate the meaning and significance of Luxembourgish idioms, one would need to learn the language and culture in depth.

- 4. "Bretzelsonndeg" is a Luxembourgish tradition that takes place on the Sunday before Lent, also known as "Quinquagesima Sunday" or "Shrove Sunday." On this day, it is customary for men to give a "Bretzel" (a type of

pretzel) to the woman they admire or love. The Bretzel is often accompanied by a small gift, such as a flower or a piece of chocolate.

The tradition of Bretzelsonndeg is believed to have originated in the Middle Ages, when pretzels were considered a symbol of good luck and prosperity. The shape of the pretzel, with its twisted form, is also thought to represent the bond of love between two people.

In modern times, Bretzelsonndeg has become a popular tradition in Luxembourg, and it is not uncommon for people to exchange Bretzels with multiple people, including friends and family members. Some people even organize Bretzelsonndeg parties or events, where they can share Bretzels and other treats with a large group of people.

If you are in Luxembourg during Bretzelsonndeg, you may want to consider participating in this tradition by giving a Bretzel to someone you care about. You can find Bretzels in many bakeries and pastry shops throughout the country, and they are often decorated with salt, poppy seeds, or sesame seeds.

In addition to giving Bretzels, you can also celebrate Bretzelsonndeg by enjoying other traditional Luxembourgish foods and drinks, such as "Grompereschelcher" (potato pancakes), "Bouneschlupp" (green bean soup), and "Fierkels" (smoked pork). Many restaurants and cafes in Luxembourg also offer special menus and promotions for Bretzelsonndeg, so be sure to check them out!

5. "Lëtzebuerg, d'kleine, aber reiche Groussherzogtum, huet eng riiche Historien a Kulture. D'Land, géint der Grenzen vun Däitschland, Frankräich a Belgiën, huet eng eisegeschichtlech Verbindung mat allen dësen Länden. D'Lëtzebuerger Sprooch, déi eng Mischung vun de Däitsch a de Franséisch Sproochen ass, ass eng wichteg Bestanddeel vun d'Lëtzebuerger Identitéit. D'Kultur vum Land ass e bësonders reich an Traditionen a Bräuchen, well eng dës méi wichtegsten ass de "Buergebrennen", déi am 31. Dezember ewéi am "Dikemesse" gefeiert gëtt. D'Lëtzebuerger Kichen ass e bësonders gudden, well eng dës méi beliebten Gerichten sinn "Judd mat Gaardebounen", "Kniddelen mat Speck" a "Gromperen zoueres". Lëtzebuerg ass och hëllefere vun eiser Natur a Landschaft, well d'Land eng hale vun Wälder, Flüssen a Fiels huet. D'Lëtzebuerger Bevëlkerung ass multikulturell a multilingual, well eng dës méi wichtegsten Minderheiten sinn d'Portugieser, d'Italiener a d'Franséiser."

(Translation: "Luxembourg, the small, but rich Grand Duchy, has a rich history and culture. The country, located at the borders of Germany, France, and Belgium, has a historical connection with all of these countries. The Luxembourgish language, which is a mixture of the German and French languages, is an important component of Luxembourgish identity. The culture of the country is particularly rich in traditions and customs, one of the most important of which is the "Buergbrennen," which is celebrated on December 31st or on "Dikemesse." Luxembourgish cuisine is particularly good, with some of the most popular dishes being "Judd mat Gaardebounen," "Kniddelen mat Speck," and "Gromperen zoueres." Luxembourg is also rich in nature and landscape, with the country having a large amount of forests, rivers, and fields. The Luxembourgish population is multicultural and multilingual, with one of the most important minorities being the Portuguese, Italians, and French.")

Task 5: Error Detection

The sentence you provided contains a few errors. Here is a corrected version:

"D'Letzebuenger Land huet 600'000 Einwohner und ass méi grouss als Däitschland."

The corrections I made are as follows:

- "Letzebuergeschen" should be "Letzebuenger" (without the "eschen" suffix)
- "Awunner" should be "Eisebenger" (the correct term for "inhabitants" in Luxembourgish)
- "2000" should be "600'000" (the correct number of inhabitants in Luxembourg)
- "grous" should be "grouss" (with an apostrophe to indicate the omission of the letter "e" before "s")

I hope this helps! Let me know if you have any other questions.

Task 6: Conversational Dialogue

"Lëtzebuerg ass e kleen, aber reicher Staat mat engem interessanten Historie a Kultur. D'Stad Lëtzebuerg, déi Hauptstodt vum Land, huet eng vill Kierchen, Schlüsselen a Museen zu besichten. E puer Woch kannsch du ech hëllefen, d'Land z'entdecken, an e puer Spaziergäנג am Fluss Alzette mécht ech mat

dir. Mir ginn och nach Lëtzebuenger Kichen probéieren, well ech glécklech sinn, datt dir d'Gerichte schmecke!"

(Translation: "Luxembourg is a small, but rich country with an interesting history and culture. The city of Luxembourg, the capital of the country, has many churches, castles, and museums to visit. In a few weeks, I can help you discover the country, and I can take you on a walk along the river Alzette. We will also try Luxembourgish cuisine, and I am happy that you will like the dishes!")

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Task 7: Luxembourgish Questions:

1. In Luxembourg, the head of government is not called the "Premierminister" (Prime Minister), but rather the "Premier" or "Première ministre" (depending on the gender of the person holding the position). The current Premier of Luxembourg is Xavier Bettel, who has been in office since December 4, 2013.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

2. In Luxembourg, the official language is Luxembourgish, which is a Germanic language that is similar to German but has its own unique grammar, vocabulary, and pronunciation. In addition to Luxembourgish, many people in Luxembourg also speak French and German, which are both considered administrative languages of the country.

In general, Luxembourgish is the most commonly spoken language in everyday conversations and in the media, while French is the language of administration and education, and German is often used in business and in the workplace.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

3. The capital of Luxembourg is the city of Luxembourg, which is also known as "Stad Lëtzebuerg" in Luxembourgish. The city of Luxembourg is located in the south of the country, near the borders of France and Germany. It is the largest city in Luxembourg and serves as the country's administrative, cultural, and economic center.

The city of Luxembourg has a rich history and is home to many historic sites, including the Bock Casemates, the Grand Ducal Palace, and the Notre-Dame Cathedral. It is also a hub for international finance, with many banks and financial institutions based in the city.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Here is a sentence in Luxembourgish about the capital of Luxembourg:

"Stad Lëtzebuerg ass d'Hauptstad vum Land a huet e bësonders interessant Historie a Kultur."

(Translation: "The city of Luxembourg is the capital of the country and has a particularly interesting history and culture.")

4. Luxembourg has a population of approximately 634,730 people (as of 2021). This makes it one of the smallest countries in Europe in terms of population, but it is also one of the most densely populated.

The majority of the population of Luxembourg is made up of Luxembourgish citizens, but there are also significant numbers of foreign residents, including people from Portugal, Italy, France, and Germany.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Here is a sentence in Luxembourgish about the population of Luxembourg:

"D'Letzebuerger Land huet 634'730 Einwohner (2021)."

(Translation: "The population of Luxembourg is approximately 634,730 (2021).")

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

5. The head of state of Luxembourg is the Grand Duke or Grand Duchess, who is the monarch of the country. The current Grand Duke of Luxembourg is Henri, who has been in office since October 7, 2000.

The Grand Duke or Grand Duchess is the chief of state and is responsible for representing the country at home and abroad, as well as for appointing the government and signing laws. The Grand Duke or Grand Duchess also has a ceremonial role and is involved in various cultural and social events.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Here is a sentence in Luxembourgish about the head of state of Luxembourg:

"D'Groussherzog oder d'Groussherzogin ass d'Wieder haut vu Lëtzebuerg."

(Translation: "The Grand Duke or Grand Duchess is the head of state of Luxembourg.")

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

6. The Grand Duke of Luxembourg in the year 2000 was Henri, who has been the Grand Duke of Luxembourg since his accession to the throne on October 7, 2000. Henri is the son of Grand Duke Jean, who abdicated the throne in favor of his son in 2000.

Grand Duke Henri is the head of state of Luxembourg and is responsible for representing the country at home and abroad, as well as for appointing the government and signing laws. He also has a ceremonial role and is involved in various cultural and social events.

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Here is a sentence in Luxembourgish about the Grand Duke of Luxembourg in the year 2000:

"D'Groussherzog vum Lëtzebuerg am Joer 2000 war Henri, déi den Tron am 7. Oktober 2000 erëm ass."

(Translation: "The Grand Duke of Luxembourg in the year 2000 was Henri, who ascended to the throne on October 7, 2000.")

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Task 8: Summarization:

"Lëtzebuerg ass e representativ Demokratie a konstitutionell Monarchie mat engem Groussherzog als Staatschef. D'Land huet e dynamesch Wirtschaft, déi zu engem vun de räichsten a räichste Länner vun der Welt gemaach huet, mat engem héichsten Bruttoinlandsprodukt pro Kapp. D'Wirtschaft konzentréiert sech op finanziell Aktivitéiten, favoriséiert durch avantagéis Besteuerung a bestëmmte Beräicher. Lëtzebuerg huet historesch eng Plaz vu grousser strategescher Wichtigkeet fir vill Muechten gemaach, an de Bau vun engem feudal Buerg am Südoste vun de Groussniederlanden, tëscht dem 15. an dem 17. Joerhonnert."

(Translation: "Luxembourg is a representative democracy and a constitutional monarchy with a Grand Duke as head of state. The country has a dynamic economy, which has made it one of the richest countries in the world, with one of the highest gross domestic products per capita. The economy is focused on financial activities, favored by advantageous taxation and certain sectors (virtual tax exemption for profits from the exploitation of patents or software). Historically, Luxembourg has had a place of great strategic importance for many countries, from its beginnings as a hinterland of the Roman Rhineland, to the construction of a feudal city on the ruins of a free Roman fortress, in a post-Carolingian context in the High Middle Ages, to its strategic location in the south-east of the Low Countries between the 15th and 17th centuries.")

I hope this helps! Let me know if you have any other questions.

(Note: I am a machine learning model and my responses may not always be 100% accurate. I always recommend double-checking my responses with a human or a reliable source.)

Task 9: Extractive question answering:

1. Here are 5 questions and answers in Luxembourgish:

- Question 1: Wat ass den Ufank vum Groussherzogtum Lëtzebuerg?
- Answer 1: Den Ufank vum Groussherzogtum Lëtzebuerg ass 1815 um Wiener Kongress.
- Question 2: Wéi gouf de Wëllem vun Oranien-Nassau zum Groussherzog vu Lëtzebuerg?
- Answer 2: De Wëllem vun Oranien-Nassau gouf als Personalunioun Groussherzog vu Lëtzebuerg.
- Question 3: Wat ass d'Groussherzogtum Lëtzebuerg an den Däitsche Bond integréiert ginn?
- Answer 3: D'Groussherzogtum Lëtzebuerg gouf an den Däitsche Bond integréiert, an d'Festung Lëtzebuerg gouf eng "Bundesfestung".
- Question 4: Warum huet de Wëllem I. sech fir "Grouss"-Herzog ze ginn?
- Answer 4: De Wëllem I. huet sech fir "Grouss"-Herzog ze ginn, well hien als Kinnek vu Lëtzebuerg wéi ee Kinnek wollt behandelt ginn.
- Question 5: Wat ass 1839 passéiert?
- Answer 5: 1839 gouf d'Groussherzogtum an zwee gedeelt: de westlechen Deel gouf als Provënz Lëtzebuerg der Belsch zougesprach, an de germanophonen Deel ass weiderhin als Groussherzogtum ginn.

2. Here are 3 questions and answers in Luxembourgish:

- Question 1: Wat ass d'Wichtegst fir e sexuellen Zesummespill?
- Answer 1: D'Wichtegst fir e sexuellen Zesummespill ass, datt ee mat vill Gefill op säi Partner ageet an datt ee verstoen, wat deem anere Pleséier mécht.
- Question 2: Warum kann d'Gréisst beim Geschlechtsverkéier eng grouss Roll spillen?

- Answer 2: D'Gréisst kann beim Geschlechtsverkéier eng grouss Roll spillen, well net all Stellung de selweschten EUUID huet. D'Positioun an d'Technik spillen och eng grouss Roll.
- Question 3: Wat hëlleft fir e richtigen Ausleben vun der Sexualitéit?
- Answer 3: Wat hëlleft fir e richtigen Ausleben vun der Sexualitéit ass, wann een op seng Partnerin an op där hir Kierpersprooch lauschtert.

3. Here are 3 questions and answers in Luxembourgish:

- Question 1: Wat heescht den neisten CD vun der Mariah Carey?
- Answer 1: Den neisten CD vun der Mariah Carey heescht "Glitter".
- Question 2: Wat ass d'Neist Auskopplung aus dem aktuellen CD vun der Mariah Carey?
- Answer 2: D'Neist Auskopplung aus dem aktuellen CD vun der Mariah Carey heescht "Never Too Far".
- Question 3: Warum ass d'Mariah Carey deen eenzëgen Artist deen all Joër eng Nummer 1 an dën Charts zë placéieren?
- Answer 3: D'Mariah Carey ass deen eenzëgen Artist deen all Joër eng Nummer 1 an dën Charts zë placéieren, well si déi meescht Nummer 1 Hits hat, wéi keen aneren weiblëchen Artist fierdrun.

Appendix C - Gemma2-9b-It model answers

Task 1: Language Understanding

Here are a few ways to paraphrase "Lëtzebuerg ass eng kleng Staat" in Luxembourgish, keeping it helpful and actionable:

- Lëtzebuerg huet eng limitéiert Fläch. (Luxembourg has a limited area.)
- D'Land Lëtzebuerg ass relativ gutt beschtëlt. (The country Luxembourg is relatively wellpopulated.)
- Wéi eng kleng Republik, huet Lëtzebuerg eng héich Bevëlkerungsdicht. (As a small republic, Luxembourg has a high population density.)

These options focus on different aspects of the original sentence and provide slightly different nuances.

Task 2: Sentiment Analysis

1. The sentiment of the text "D'Wirtschaft vum Land ass momentan schlecht" is negative.

The phrase "ass momentan schlecht" directly translates to "is currently bad," clearly indicating a negative sentiment towards the state of the country's economy.

2. The sentiment of the text "Ech hunn dech gär, mee mir passen net zesummen" is negative.

While "Ech hunn dech gär" means "I like you," the phrase "mee mir passen net zesummen" translates to "but we don't fit together." This clearly indicates a sense of incompatibility and ultimately a negative sentiment towards the relationship.

3. The sentiment of the text "Ech hunn haut 300€ am Lotto gewonnen" is positive.

Winning money in the lottery is generally considered a positive event, and the phrase expresses joy and excitement.

4. The sentiment of the text "Ech hunn haut 300€ op der Strooss verluer" is negative.

Losing money is generally a negative experience, and the phrase expresses a sense of loss and likely frustration.

5. The sentiment of the text "Ech hunn haut mäin Brudder getraff" is negative. "Getraff" in this context means "hit" or "attacked," implying a physical altercation. Physical violence is generally considered a negative action.

Task 3: Translation

1. Here are a couple of ways to translate "The weather is nice today" into Luxembourgish:

- D'Wéiwer ass haut gutt. (This is a more common and straightforward translation.)
- Et ass e schönen Dag. (This emphasizes the pleasantness of the day.)

Let me know if you'd like to see other variations!

2. The translation of "The European Union has its headquarters in Brussels" in Luxembourgish is:

D'Europäesch Unioun huet hir Hauptsitz an Brüssel.

Let me know if you have any other sentences you'd like me to translate!

3. The English translation of "Ech ginn an de Park" is:

I am going to the park.

Let me know if you have any other Luxembourgish phrases you'd like translated!

4. Here are a couple of ways to say "I'm going to the park" in Luxembourgish:

- Ech ginn an de Park. (This is the most common and straightforward translation.)
- Ech fahre/geet an de Park. (This emphasizes the act of traveling to the park.)

Let me know if you have any other phrases you'd like translated!

5. Here are a couple of ways to translate "Luxembourg is a small country in Europe" into Luxembourgish:

- Lëtzebuerg ass e klengen Land an Europa. (This is a more literal translation.)

- Lëtzebuerg ass e béidlech Land an Europa. (This emphasizes the small size of Luxembourg.) Let me know if you have any other phrases you'd like translated!

Task 4: Text Generation

1. Lëtzebuergesch Kultur ass eng faszinéierend Mixtur aus traditioneller Geméinlichkeit an moderner Offenheit, geprägt vum Schmelztiegel vun verschiddene Sprochen an kulturellen Einflüssen.
2. Am Moseltal, wou d'Wijn so sweet, D'Steeën stonn, an d'Sonn so heit. D'Wälle flieë, d'Bächer glei, An d'Natur ass e Traum, fir all' Zee.

De Schëff gëtt, am Duerf so klein, D'Leit sinn frëndlech, an d'Wijn so fein. Am Moseltal, wou d'Leed so schön, D'Hëllef vun der Natur, ass e Geschicht ze ginn.

Translation:

In the Moselle Valley, where the wine is so sweet, The stones stand, and the sun so bright. The waves flow, the rivers gleam, And nature is a dream, for all to see.

The boat goes, in the village so small, The people are friendly, and the wine so fine. In the Moselle Valley, where the songs are so beautiful, The help of nature, is a story to be told.

3. Here are five Luxembourgish idioms with their approximate English translations:
 - a) "Ech hu sech eng Schlass gepackt." - I've gotten myself into a mess.
 - b) "D'Schëff ass op d'Wäer." - Things are going smoothly.
 - c) "Et kritt eng Kréiz." - It's getting complicated.
 - d) "Hien huet sech eng Flënzer gepackt." - He's gotten himself into a difficult situation.
 - e) "Et ass net alles Gold, wat glitt." - All that glitters is not gold.

Let me know if you'd like to know more about Luxembourgish idioms!

4. Bretzelsonndeg ass e traditionelle Feierdag an Lëtzebuerg, dee op d'Sonndeg nom Pfingstsonndeg gefeiert gëtt.

Et ass e Dag fir Famill an Frënn zesumme ze kënnen, déi leckeren Brezel ze genéisze, déi meeschtens mat Senf servéiert ginn, an d'Ënnergang vum Fruehjahr ze fieren.

D'Tradition ass schonn vill Joer a Lëtzebuerg etabléiert a beziechent d'Geméinlechkeet an d'Gemeinschaftsverbundenheet vun de Lëtzebuerger.

5. Lëtzebuerg, e Land am Zentrum vun Europa, huet eng reiche Geschicht an eng faszinéierend Kultur. Als eng vun de klengsten Länner vun der EU, ass et bekannt fir seng wunderschöne Landschaften, seng historische Schloesser an seng kosmopolitisch Geméin. Déi lëtzebuergesch Sprooch, eng romanesch Sprooch, ass eng Mixtur aus Franséisch, Däitsch an Italienesch. D'Lëtzebuerger Kultur ass geprägt vum Schmelztiegel vun verschiddene Einflüssen, déi sech an der Kënscht, der Musek an der Gastronomie widerspiegeln. D'Land huet eng héich Lebensqualitéit a bekannt fir seng Toleranz an d'Geméinlechkeet.

Task 6: Conversational Dialogue

Lëtzebuerg ass e charmantes Land am Zentrum vun Europa, bekannt fir seng wunderschöne Natur, seng historische Schloesser an seng kosmopolitisch Geméin. D'Land huet eng reiche Kultur, déi sech an der Kënscht, der Musek an der Gastronomie widerspiegelt. Komm an Lëtzebuerg a entdeckt d'Geméinlechkeet, d'Toleranz an d'héich Lebensqualitéit!

Task 7: Luxembourgish Questions:

1. De Premierminister vu Lëtzebuerg ass Xavier Bettel.
2. D'chính thục Sprooch vu Lëtzebuerg ass Lëtzebuergesch, eng romanesch Sprooch.

Allerdings gëtt och Franséisch an Däitsch ganz wéichtg geschwat, an et gëtt eng meng vun aner Sproochen vun Immigréinte gesprochene.

3. D'Hauptstad vu Lëtzebuerg ass Lëtzebuerg (oder "Luxembourg" op Däitsch an Franséisch).
4. Lëtzebuerg huet ongeféier 640.000 Awunner.
5. Als KI kann ech keng Echtzeit-Informatioun, wéi d'Wetter, bereitstellen. Fir d'aktuell Wetterbericht fir Lëtzebuerg ze konsultieren, kann ech dir eng Wetter-App oder -Websäit recommendéieren.

6. De Groussherzog vu Lëtzebuerg am Joer 2000 war Jean.

Task 8: Summarization:

Lëtzebuerg ass eng Demokratie mat engem Groussherzog als Staatschef. D'Wirtschaft ass finanzgeprägt a deemno een vun de räichsten Länner. D'Land huet eng historesch strategesch Plaz an Europa.

Task 9: Extractive question answering:

1. Here are 5 questions and answers in Luxembourgish based on the text:

- 1. Wann gouf d'Herzogtum Lëtzebuerg nees begrënnt?
- A: D'Herzogtum Lëtzebuerg gouf 1815 um Wiener Kongress nees begrënnt.
- 2. Wéi gouf d'Gebiet vum neie Lëtzebuerger Herzogtums definéiert?
- A: Et gouf aus den al lëtzebuergesch Gebidder ëstlech vun der Our, mat Ausnam vu Veianen, an Deeler vum Departement des Forêts an Deeler vum Departement Sambre-et-Meuse geformt.
- 3. Wem gouf d'Lëtzebuerger Thron zugewies?
- A: De Wëllem vun Oranien-Nassau, dee grad kuerz virun als hollännesche Kinnek Wëllem I. unerkannt gi war, gouf Groussherzog vu Lëtzebuerg.
- 4. Wéi ass d'Bezeichnung "Groussherzogtum" entstanen?
- A: Wëllem I. huet fir "Grouss"-Herzog ze gin, fir als Kinnek behandelt ze ginn, wat en als "einfachen" Herzog vu Lëtzebuerg net hätt kënne sinn.
- 5. Wann gouf d'Groussherzogtum Lëtzebuerg an zwee Deel opgedeelt?
- A: Den 19. Abrëll 1839 gouf d'Groussherzogtum an zwee Deel opgedeelt: de westlechen Deel gouf der Belsch zugewies, an de germanophonen Deel bleift als Groussherzogtum.

2. Here are 3 questions and answers in Luxembourgish based on the text:

- 1. Wat ass wichteg fir en erfuhlleche Geschlechtsverkéier?
- A: D'Zesummespill vu béide Partner ass wichteg. D'Gréisst ass net d'wichtigste Faktor
- 2. Wéi kann d'Mythos iwwert d'Gréisst Männer ënner Drock setze?

- A: D'Mythos kann Männer ënner Drock setze, fir hir Sexualitéit richteg auszeliewen an ze genéissen. Dëse Drock kann och zu Operatiounen fir Verlängerung leiten.
- 3. Wat kann Männer hëllefen, fir hir Sexualitéit besser ze genéissen?
- A: D'Partnerin an hir Kierpersprooch lauschten ass keng Verlängerung. D'Kommunikatioun an d'Zesummespill sinn déi wichtegsten Faktoren.

3. Here are 3 questions and answers in Luxembourgish based on the text:

- 1. Wat ass de Titel vum neie Album vun der Mariah Carey?
- A: De Titel vum neie Album vun der Mariah Carey ass "Glitter".
- 2. Wéi heescht d'neist Auskopplung aus dem Album "Glitter"?
- A: D'neist Auskopplung aus dem Album "Glitter" heescht "Never Too Far"
- 3. Wann huet d'Mariah Carey hir éischt Filmrolle eegenduecht?
- A: D'Mariah Carey huet hir éischt Filmrolle am Film "Glitter" eegenduecht.

Appendix D - Llama 3.1-405b-instruct model answers

Task 1: Language Understanding

Lëtzebuerg ass e klengt Land mat enger eegener Regierung.

(Note: I've added more context to the paraphrased sentence, but a more direct paraphrase would be "Lëtzebuerg ass e klengt Land")

Task 2: Sentiment Analysis

1. Negative
2. Neutral/Bittersweet

(Note: The text expresses a mix of positive and negative emotions, as the speaker says they like the person, but also states that they are not compatible. This ambivalence makes the overall sentiment neutral or bittersweet)

3. Positive
4. Negative
5. Positive

(Note: The text implies a neutral or even positive sentiment, as meeting a family member is typically a pleasant experience. However, without more context, it's difficult to determine the exact sentiment. Assuming a typical familial relationship, I've labeled it as positive)

Task 3: Translation

1. D'Wieder ass haut schéin
2. D'Europäesch Unioun huet hire Sëtz zu Bréissel.
3. I'm going to the park.
4. Ech ginn an de Park.
5. Lëtzebuerg ass e klengt Land an Europa.

Task 4: Text Generation

1. D'Lëtzebuergesch Kultur ass eng räich a vielfälteg Mëschung vu franséischen, däitschen an aneren europäeschen Aflëss, déi sech a Konscht, Literatur, Musek a Bräich widerspiegelt. Si ass geprägt vu senger Geschicht, senger geographescher Lag an der Zweisproochegkeet vu sengen Awunner.

2. D'Musel, déi schéin, Lët sech duerch d'Hiwwelen zeien, Wéi e silwere Band, Déi duerch d'Wéngerten a Blummen zeiert.

D'Sonn schéngt op d'Wasser, Déi sech a Wellen a Kreesen beweegt, D'Vigelcher sangen hiren Déif, An d'Luucht ass erfüllt vu Freed a Séilegkeet.

D'Musel, déi wunderschéin, Ass e Paradies op Äerden, Wou een sech erholen a friddele kann, An d'Schéinheet vun der Natur genéissen kann.

3. Here are five Luxembourgish idioms:

- a) "Ech hunn d'Schnur verluer" - This means "I've lost the thread" or "I've lost my train of thought".
- b) "Dat ass keng Kaffeesatz" - This means "That's no coffee grounds" or "That's nothing to worry about".
- c) "Ech si mat de Féischt am Dreck" - This means "I'm up to my neck in trouble" or "I'm in a difficult situation".
- d) "Dat ass eppes fir d'Aen" - This means "That's something for the eyes" or "That's a feast for the eyes".
- e) "Ech hunn e Knascht am Kapp" - This means "I have a knot in my head" or "I have a problem that I'm struggling to solve".

Note: Luxembourgish idioms often use metaphors and wordplay, and may not translate directly into other languages.

4. De Bretzelsonndeg ass e lëtzebuergesche Brauch, deen all Joer op Pällemsonndeg gefeiert gëtt. Op dësem Dag ginn d'Männercher Bretzele fir hir Fraen oder Frëndinnen. D'Bretzele sinn aus Bounoder Weessmehl gemaach a gi mat Salz oder Zucker bestreut.

D'Traditioun geet zeréck op d'Zäit, wou d'Männercher op Pällemsonndeg hir Fraen oder Frëndinnen mat Bretzele begéinten, fir hinnen ze beweisen, datt si sech ëm si këmmern. Hautdesdaags gëtt de Bretzelsonndeg nach ëmmer gefeiert, awer et ass méi eng Geleeënheet fir sech mat Frënn a Famill ze treffen an zesummen Bretzele ze genéissen.

D'Bretzele ginn traditionell vun de Männercher selwer gemaach oder gekaaft, an da gi si den Fraen oder Frëndinnen iwwerreecht. Et ass e schéinen Dag, fir sech mat deene Leit, déi een gäre huet, ze treffen an zesummen ze feieren.

5. Lëtzebuerg huet eng räich a vielfälteg Geschicht, déi vu réimeschen a fränkeschen Aflëss geprägt ass. Déi éischt Awunner vu Lëtzebuerg waren d'Kelten, déi sech do ëm 500 v. Chr. néiergelooss hunn. Duerno hunn d'Réimer d'Regioun eruewert an do hir Spueren hannerlooss. Am Mëttelalter huet Lëtzebuerg zu de wichtegste Fürstentümer vu Westeuropa gezielt, an d'Grouss vu Lëtzebuerg hunn eng wichteg Roll an der europäescher Politik gespillt. Hautdesdaags ass Lëtzebuerg e modernen a vielfältegen Staat, deen eng räich Kultur a Geschicht huet. Déi lëtzebuergesch Sprooch, déi zu de westzentrale germanesche Sproochen gezielt gëtt, ass eng wichteg Deel vu Lëtzebuergs Identitéit, an d'Awunner vu Lëtzebuerg si stolz op hir Geschicht a Kultur.

Task 5: Error Detection

There are two errors in the sentence:

1. "Den Letzebuergeschen Land" should be "D'Land Lëtzebuerg" or simply "Lëtzebuerg". The correct way to refer to the country is by using the definite article "d'" or no article at all.
2. The statement "huet 2000 Awunner" is incorrect. According to the Statec Institute, Luxembourg has a population of around 590,000 people (as of 2020). Additionally, the statement "ass méi grouss als Daitschland" is also incorrect, as Luxembourg is actually one of the smallest countries in Europe, while Germany is one of the largest.

Corrected sentence: "Lëtzebuerg huet ronn 590.000 Awunner an ass ee vun de klengste Länner an Europa."

Translation: "Luxembourg has around 590,000 inhabitants and is one of the smallest countries in Europe."

Task 6: Conversational Dialogue

Lëtzebuerg ass e klengt Land mat enger räicher Geschicht a Kultur. Et ass bekannt fir seng schéi Landschaften, wéi d'Äerzgebierg an d'Musel, an och fir seng al Stied, wéi d'Stad Lëtzebuerg, déi op enger schéiner Kopp läit. Lëtzebuerg huet och eng wichteg Roll an der europäescher Politik gespillt an ass hautdesdaags e wichteg Zentrum fir Finanzen an Technologie.

Task 7: Luxembourgish Questions:

1. De Premierminister vu Lëtzebuerg ass de Xavier Bettel. Hien ass zanter dem 4. Dezember 2013 Premierminister a Minister fir Staats- a Kommunikatiounsaffären.
2. Zu Lëtzebuerg gi véier offiziell Sproochen geschwat:
 - a) Lëtzebuergesch (national Sprooch)
 - b) Franséisch
 - c) Däitsch
 - d) Englesch (als véiert offiziell Sprooch zanter 1984)

D'Lëtzebuergesch ass d'national Sprooch a gëtt vu ronn 70 Mammesprooch geschwat. Franséisch a Däitsch sinn och weit verbreet a ginn als zweet Sprooch geschwat. Englesch ass als véiert offiziell Sprooch an der Verwaltung, der Wirtschaft an der Bildung verbreet.

3. D'Hauptstad vu Lëtzebuerg ass Lëtzebuerg-Stad (oder einfach Lëtzebuerg). Si ass déi gréisst Stad vu Lëtzebuerg a läit am Zentrum vum Land, op der Kruzung vun den dräi Flëss Alzette, Pétrusse a Sauer. Lëtzebuerg-Stad huet ronn 120.000 Awunner an ass déi wirtschaftlech, kulturell a politesch Zentral vu Lëtzebuerg.
4. Lëtzebuerg huet ronn 590.000 Awunner (Stand 2020). D'Bevëlkerung vu Lëtzebuerg ass eng vun de klengsten an Europa, awer si ass och eng vun de räichsten a méi diversifiéierten.
5. Dat kann ech leider net soen, well ech net an der Lag sinn, aktuell Wiederdaten ze liesen. Mä ech kann lech awer suggeréieren, op d'Internetseite vum Lëtzebuerger Wiederamt (Administration de la navigation aérienne) oder op aner Wieder-Websäiten ze goen, fir déi aktuell Wiederdaten ze kucken.
6. 6 De Groussherzog vu Lëtzebuerg am Joer 2000 war de Jean vu Lëtzebuerg (Jean, Grand-Duc de Luxembourg). Hien ass den 12. November 1921 gebuer ginn a war vum 12. November 1964 bis den 7. Oktober 2000 Groussherzog vu Lëtzebuerg. Den 7. Oktober 2000 huet hien ofgedankt an säi Jong, de Henri, gouf neie Groussherzog vu Lëtzebuerg.

Lëtzebuerg ass eng representativ Demokratie a konstitutionell Monarchie mat engem Groussherzog als Staatschef. D'Land huet eng dynamesch Wirtschaft a gehéiert zu de räichsten Länner vun der Welt. D'Wirtschaft

konzentréiert sech haaptsächlech op finanziell Aktivitéiten a profitéiert vu avantagéis Besteierung.

Task 9: Extractive question answering:

1. Here are 5 questions and answers in Luxembourgish based on the text:

- Q1: Wat ass um Wiener Kongress 1815 geschitt? A1: Um Wiener Kongress 1815 gouf dat viregt Herzogtum Lëtzebuerg vun den europäesche Groussmuechten net "restauréiert".
- Q2: Wéi gouf de Wëllem vun Oranien-Nassau 1815 ernannt? A2: De Wëllem vun Oranien-Nassau gouf 1815 a Personalunioun Groussherzog vu Lëtzebuerg ernannt.
- Q3: Wat ass d'Festung Lëtzebuerg 1815 ginn? A3: D'Festung Lëtzebuerg gouf 1815 eng "Bundesfestung", wat dem preisesche Kinnekräich d'Garnisounsrecht erméiglecht huet.
- Q4: Wéi ass Lëtzebuerg 1839 gedeelt ginn? A4: Lëtzebuerg gouf 1839 um Traité vu London an zwee gedeelt: de westlechen Deel gouf der Belsch zougesprach, an de germanophonen Deel ass weiderhin als Groussherzogtum fir d'éischt Kéier een eegestännege Staat ginn.
- Q5: Wéi gouf de Wëllem I. als Groussherzog vu Lëtzebuerg behandelt? A5: De Wëllem I. gouf als Groussherzog vu Lëtzebuerg behandelt wéi ee Kinnek, an huet drop gehale fir "Grouss"-Herzog ze ginn, also eng "Kinneklech" Houheet.

2. Here are 3 questions and answers in Luxembourgish based on the article:

- Q1: Wat ass wichteg beim Geschlechtsverkéier? A1: Dat wat zielt ass d'Zesummespill vu béide Partner, an et kënnt net op d'Gréisst vum Mann sengem beschte Stéck un.
- Q2: Wéi kann ee seng Sexualitéit richteg auszeliewen an ze genéissen? A2: Ee kann ee seng Sexualitéit richteg auszeliewen an ze genéissen, wann ee op seng Partnerin an op där hir Kierpersprooch lauschtert, an net ënner Drock geréit.
- Q3: Wat ännert eng Verlängerung vum Mann sengem beschte Stéck? A3: Eng Verlängerung ännert net onbedéngt eppes, well déi näischt un der Sensibilitéit géigeniwwer der Partnerin ännert.

3. Here are 3 questions and answers in Luxembourgish based on the article:

- Q1: Wat ass den Titel vum neisten CD vun der Mariah Carey? A1: Den Titel vum neisten CD vun der Mariah Carey ass "Glitter".
- Q2: Wat ass d'Spezialitéit vun der Mariah Carey hirer Stëmm? A2: D'Mariah Carey hir Stëmm geet iwwert 8 Octaven, an si kann verschidde Facetten vu senger Stëmm zë weisen.
- Q3: Wat ass d'Zil vum Album "Glitter"? A3: D'Zil vum Album "Glitter" ass, eng musekalësch Rees duërch dëm Billie säi Geescht duerstellen, an de Sound vun dën 80er ass op dësem Album omniprésent.