Journal of the Brazilian Computer Society, 2025, 31:1, doi: 10.5753/jbcs.2025.5795

Domain Learning from Data for Large Language Model Translation and Adaptation

René Vieira Santin [**Universidade de São Paulo** | ***renevs@usp.br***]
Ricardo Marcondes Marcacini [**Universidade de São Paulo** | ***ricardo.marcacini@usp.br***]
Solange Oliveira Rezende [**Universidade de São Paulo** | ***solange@icmc.usp.br***]

Instituto De Ciências Matemáticas e de Computação (ICMC), USP - São Carlos, Av. Trab. São Carlense, 400, Parque Arnold Schimidt, São Carlos, SP, 13566-590, Brazil.

Received: 31 March 2025 • **Accepted:** 30 July 2025 • **Published:** 21 October 2025

Abstract Large Language Models (LLMs) have improved multilingual translation and adaptation, particularly for languages like Portuguese; however, they often fail to produce outputs that accurately reflect the linguistic, stylistic, and topical characteristics expected in real-world scenarios. This paper addresses the challenge of adapting texts to specific domains and audiences by moving beyond direct translation to include variations in genre and topic. We propose a method for learning domain representation vectors through prompt tuning, allowing LLMs to generate text that matches the communicative norms of a target domain or user profile (e.g., legal discourse, informal speech, or social media posts) or even topics. In contrast to most domain adaptation approaches that focus solely on translation, our method supports broader text adaptation and can be applied to multiple tasks. We demonstrate the effectiveness of our approach using two Portuguese datasets—a newly compiled corpus of video game discussions and a financial tweet corpus—and evaluate the results with respect to linguistic variation. Our main contributions include: (i) a method for learning reusable domain vectors to support prompt-based adaptation; (ii) application to translation and broader text adaptation tasks; and (iii) the release of a new domain-specific dataset in Portuguese.

Keywords: Translation, Domain Adaptation, Portuguese, Large Language Model, LLM, Prompt

1 Introduction

Large Language Models (LLMs) trained on multilingual corpora have become widely explored tools for adaptation and translation [Mann *et al*., 2020; Zhang *et al*., 2023a]. However, adapting and translating a foreign language into Portuguese may be subject to domain-specific variations. For example, consider the English sentence "The house burglar from the region has been arrested" (see **Table 1**). A direct translation into Portuguese would be "*O ladrão de casas da região foi preso*". However, a lawyer might translate the sentence as "*O indivíduo suspeito de cometer furtos em residências na região foi detido pelas autoridades competentes*" (The individual suspected of committing burglaries in the area has been detained by the competent authorities). A child might say: "*O moço que roubava as casas foi pego*" (The guy who took things from houses got caught).

Several factors influence domain variations, including the medium of communication. A text written by a financial agent on X (formerly Twitter) differs entirely from the same agent's text in a central bank report. Accurately determining the domain of a text remains a challenging task due to overlapping linguistic characteristics. For instance, a group of teenagers discussing video games on Reddit uses a precise vocabulary, making it challenging to express this domain simply and accurately using words.

In another context, consider a scenario in which it is desired to train a machine learning model for the sentiment analysis task in a financial domain in Portuguese, where only a limited amount of labeled data is available. Suppose there is an annotated dataset from a different domain and language, such as Laptop reviews in English, which could be leveraged for this task. In this case, to mitigate domain shift, one may instruct a large language model (LLM) to translate and adapt the Laptop-domain sentences to resemble those from the financial domain. However, even after this adaptation, the resulting data may still differ significantly from the target domain in terms of style, formality, vocabulary, syntactic structure, and other linguistic characteristics. As a result, domain shift may persist and hinder the model's performance during training.

Thus, the search for exact domain representations from a text dataset brings several benefits, some of which are listed below:

- A text in English or another language can be adapted into Portuguese for a specific agent's language in a particular medium, such as children from a given region.
- A marketing campaign for an international company can convey its message using the precise language of the target audience.
- Labeled or unlabeled data used for machine learning in one domain and language can be translated and adapted into data from another language and domain. For example, an English text labeled for sentiment analysis of restaurants can be adapted for a Portuguese-language computer domain.

This work introduces a method for obtaining domain representation vector(s) for the target data, which can be leveraged in domain translations and/or adaptations using the *prompt tuning* technique [Lester *et al*., 2021]. Prompt tuning is based

Table 1. Sentences spoken by different agents

Agent	Sentence spoken
Direct Translation	*O ladrão de casas da região foi preso.* (The house burglar from the region was arrested.)
Lawyer	*O indivíduo suspeito de cometer furtos em residências na região foi detido pelas autoridades competentes.* (The individual suspected of committing burglaries in the area has been detained by the competent authorities.)
Child	*O moço que roubava as casas foi pego.* (The man who was stealing houses got caught.)

on the premise that the input prompts for language models are composed of learnable vectors, a concept known as *soft prompting*. In the proposed approach, all prompt vectors remain fixed, except for those representing the language and domain, which are optimized during training. A two-step training process is employed to achieve this. A Large Language Model (LLM) converts target domain sentences to arbitrary domains and languages in the first step. In the second step, an LLM, with a prompt where only the vectors representing the language and domain should be learned, converts the sentences back to the target language and domain, penalizing divergent sentences. After training, the LLM uses the learned vectors to translate other data sources. By selectively learning domain-specific representations, the proposed method enables the reuse of these vectors across different prompts, such as those employed in translating labeled datasets.

The study used two datasets as targets for domain and language learning. The first dataset consists of sentences in Portuguese from discussions about console and computer games. This dataset was created from the gamesEcultura group on Reddit[1] and is the first contribution of this article. The second dataset is a corpus of stock market tweets [Zerbinati *et al*., 2024], also in Portuguese. This article describes a study on applying this model to various datasets, comparing it with translation into the supposed domain in Portuguese.

A significant difference from the works on domain adaptation for neural machine translation (DA-NMT) [Saunders, 2022; Chu and Wang, 2018] lies in the broader scope of our approach. While most DA-NMT studies focus on building translation models that perform effectively in a target domain different from the source, our objective is to learn domain representation vectors that support multiple tasks. These tasks include not only domain-adapted translation but also full-sentence adaptations—addressing changes in topic and vocabulary—for use in downstream tasks. For example, a named entity recognition model trained in the laptop domain can be enhanced by incorporating sentences adapted from the restaurant domain using our proposed method.

Finally, when evaluating the adaptation and translation of a text, it is common to use metrics that compare the presence of sequences of words in a translated text. However, such metrics become unfeasible for domain conversion analysis. Therefore, in this article, we chose to perform quantitative analysis using distance metrics between distributions, including semantic and syntactic elements, vocabulary, and the standard deviation of the meaning of the generated sentences, alongside a qualitative analysis performed with a Large Language Model (LLM).

The main contributions of this work are:

- The release of an unlabeled dataset of game discussion sentences in Portuguese.
- A methodology to learn a domain representation for context-aware translations and/or adaptations to be used in LLM prompts. This representation can be used for translations and domain adaptations, including labeled data for supervised machine learning.
- A comparison of the application of trained prompts on different datasets.

The rest of this article is organized as follows: **Section 2** presents the related work. **Section 3** provides an overview of the domain concept used in this article. **Section 4** formulates the problem. The methodology employed is described in **Section 5**. **Section 6** outlines the experiments conducted. **Section 7** presents the results. **Section 8** provides a discussion of the obtained results. The conclusion is presented in **Section 9**.

2 Related Work

The proposed work is related to the field of *Domain Adaptation for Neural Machine Translation* (DA-NMT) [Saunders, 2022; Chu and Wang, 2018], although the proposed is not focused on pure translation. DA-NMT seeks models capable of performing translations across different domains. In this context, one common challenge in this field is overfitting. Many of these models predict the next token based on input data, and often, these models become dominated by more common words when fine-tuning the models on training data. Techniques such as regularization with a uniform word distribution [Szegedy *et al*., 2015] or a teacher-student model [Hinton *et al*., 2015] may be considered.

DA-NMT research can be divided into four groups: *Data-Centric*, *Architecture-Centric Adaptation*, *Training Schemes for Adaptation*, and *Inference Schemes for Adaptation* [Saunders, 2022]. Data-centric solutions aim to generate or select data according to the target domain for model training. Li *et al*. [2018] use a subset of test sentences to select training sentences and fine-tune the pre-trained model. The model proposed by Cai *et al*. [2021] is trained in two stages: in the first stage, the model selects a relevant set of target-domain sentences for a given source sentence based on a learned function; in the second stage, the original sentence and the selected sentences are used to predict the translation. Koehn and Senellart [2010] remove source-domain data that could

[1] gamesEcultura - https://www.reddit.com/search/?q=gamesEcultura

introduce noise into the training process. Some data-centric models use back-translation; they employ a model that translates target data back to the source, generating synthetic data [Sennrich *et al*., 2016].

The *Architecture-Centric Adaptation* group explores model modifications that enable them to function across different languages and domains, which can occur in various ways. For instance, Stergiadis *et al*. [2021] pass tags to differentiate domains and the other input data to the model. You *et al*. [2024] developed a non-autoregressive NMT model based on k-Nearest Neighbors (k-NN) for domain adaptation.

The *Training Schemes for Adaptation* group focuses on solutions related to training. For example, Liu *et al*. [2024] proposed the iterative dual domain adaptation model for neural machine translation. This solution follows the premise that translation data is available from two domains. So, two models form the solution, one for each domain. Knowledge transfer occurs between them during the training, reusing the model weights and aligning the distributions.

Although these categories provide a useful taxonomy for DA-NMT systems, the approach proposed in this work does not fit directly into any of them. This is because our method does not aim to translate between domains while preserving the original meaning. Instead, it focuses on learning a domain embedding that can be used to adapt texts to match the characteristics of a target domain, which may include changes in genre and even topic. As such, our method addresses a broader task than typical DA-NMT approaches.

Finally, the *Inference Schemes for Adaptation* topic encompasses works where domain adaptation occurs after model training and seeks to manipulate inference based on the domain. An example in this topic is the work of Liu *et al*. [2020]. They use sentence similarity and the corpora of trained models to distribute the weights between inference from a generic model and models trained on different corpora.

Considering the techniques employed in these tasks, the emergence of learning models using the Transformer architecture [Vaswani *et al*., 2023], trained on large volumes of text, has brought a considerable performance gain to natural language processing tasks. Applying knowledge transfer to these models through fine-tuning has enabled tasks such as translation, summarization, and question answering to achieve near-state-of-the-art performance [Devlin *et al*., 2019; Radford *et al*., 2019].

In this same vein, general-purpose text-to-text models emerged, such as T5 [Raffel *et al*., 2023]. This model was capable of performing multiple tasks depending on the input prompt. These models evolved into the billion-parameter models we see nowadays, such as GPT-3 [Mann *et al*., 2020], GPT-4 [OpenAI *et al*., 2024], and LLAMA [Touvron *et al*., 2023]. However, fine-tuning used in the early models became impractical for standard user machines.

Rather than adjusting all the model weights, techniques have been developed that modify only a small number of parameters, with *prompt tuning* being one of them [Lester *et al*., 2021]. In this approach, the pre-trained model is kept frozen, and trainable continuous vectors — known as soft prompts — are learned and prepended to the input embeddings. These vectors are updated via gradient descent during backpropagation. Similarly, a solution can freeze part of the prompt, adjusting only specific vectors. This approach is employed in the proposal of this paper.

3 Preliminaries

The term *domain* has several definitions, with the most common being a corpus from a specific source which can differ from other domains in topic, genre, style, level of formality, etc [Koehn and Knowles, 2017]. However, Saunders [2022] points out some problems with the definition "*a corpus from a specific source*", often referred to as *provenance*:

i. The provenance of the test data is not always known;
ii. In contrast to the provenance, a sentence or text can be enough to identify the genre or topic.
iii. Topic and genre can overlap, while provenance is a discrete label for a dataset.

Instead of using provenance to describe a domain, we can analyze it from several perspectives. These attributes, which allow us to analyze whether a text aligns with a domain, include the following:

i. **topic** is the subject of a text, such as financial, medical, computer games, etc. Each topic has a vocabulary distribution, and a text can be viewed as a mixture of multiple topics [Saunders, 2022].
ii. **genre** is a macro concept that incorporates function, register, syntax, and style [Santini, 2004; Saunders, 2022]. For instance, an e-mail, a product description, and a product advertisement can have different genres for the same topics, even if they may share the same vocabulary and entity names. Automatic genre detection in text can analyze the relative frequency of different syntactic categories, the use of particular characters such as punctuation marks, and sentence length distribution [Saunders, 2022].

For a microanalysis in genre, we will use the following definitions [Santini, 2004]:

i. **function**: texts can be grouped by the communicative purpose. In other words, the function of a text may be to inform, persuade, narrate, instruct, etc.
ii. **register**: refers to the conventionalized, functional configuration of language tied to certain broad social situations. For example, a legal discourse has a technical vocabulary, a rigid structure, and an impersonal tone.
iii. **syntax**: the grammatical structures that are typically used in a genre. For example, short sentences with a predominance of active voice.
iv. **style**: refers to an individual distinct use of language.

In this study, we approach the notion of domain along two complementary axes. For the adaptation process, we adopt a provenance-based view: adapted sentences should appear as if they naturally belong to the target data source. For the analysis and evaluation, however, we rely on topic and genre as linguistic and semantic proxies to assess how well domain-specific characteristics are either preserved or transformed.

4 Problem Formulation

The problem addressed is the search for a domain embedding to be used in translation and adaptation of sentences into another language and domain. Specifically, we aim to generate sentences in a target domain in Portuguese adapted from a set of sentences in source domains in other languages, such that the generated sentences are similar to those in the target domain. We adopt the concept of domain provenance in this context. Ideally, the generated sentence should remain close to the original in meaning and structure, while being adapted **only as much as necessary** to conform to the characteristics of the new domain. Consequently, it should be difficult to determine whether a sentence originated from the target dataset or is an adapted version from a source domain.

To achieve this, the sentences must be adapted in terms of topic and genre, including function, register, syntax, and style. However, if the target domain doesn´t have a very deterministic topic, genre, or some attribute of genre, the original phrase must preserve as many of these characteristics as possible. For instance, if the target domain does not present sentences with a unique function, the function from the original sentence must be preserved. We hope the sentence preserves at least their sentiments and pragmatics. Although in extreme cases, if the destiny domain has only negative or positive sentences, the phrase should be adapted as well.

Let D_t denote the target domain, composed of sentences in Portuguese. Let $D_s^{(k)}$ be a dataset consisting of sentences in a different language and domain, where k indexes the k-th dataset among a set of K source datasets. Our goal is to learn a domain embedding w for a transformation function $f(w, x)$ such that, for the adapted dataset $D_s^{'} = \{y = f(w, x) | x \in D_s^{(k)}\}$, the distance between the distribution of $D_s^{'}$ and that of D_t is minimized. In other words, the probability of a sentence being found in any of the domains becomes very close. At the same time, the output y should preserve, as much as possible, the semantic and structural properties of the original input x, including topic, genre, pragmatic, etc.

5 Methodology

The proposed method relies solely on target data during training, which can consist of sentences in any language and domain, such as those from *Games in Portuguese*. The approach is based on partial learning of a prompt for an LLM and is conducted in two stages, as illustrated in **Figure 1**. In the first stage, an LLM with a fixed prompt — "*Adapt and translate the above sentence into the domain and language: [Domain/Language(1)]*" — adapts and translates the target data into a different domain and language, generating x'. The *[Domain/Language(1)]* is a hyperparameter of the model, such as *Movie in German*. Thus, the sentences x' maintain a relationship with x. Several alternatives for this hyperparameter were explored (**Section 6**), such as *Celebrity in German* and *Sport in English*. The adapted and translated sentences compose the *intermediate domain*.

Preliminary empirical tests were conducted to ensure that the small LLM produced only the adapted sentence in this first step, without extra commentary (e.g., "Here is the adapted sentence:"). Although we did not perform extensive prompt engineering, we refined the prompts through minor adjustments until the desired behavior was achieved.

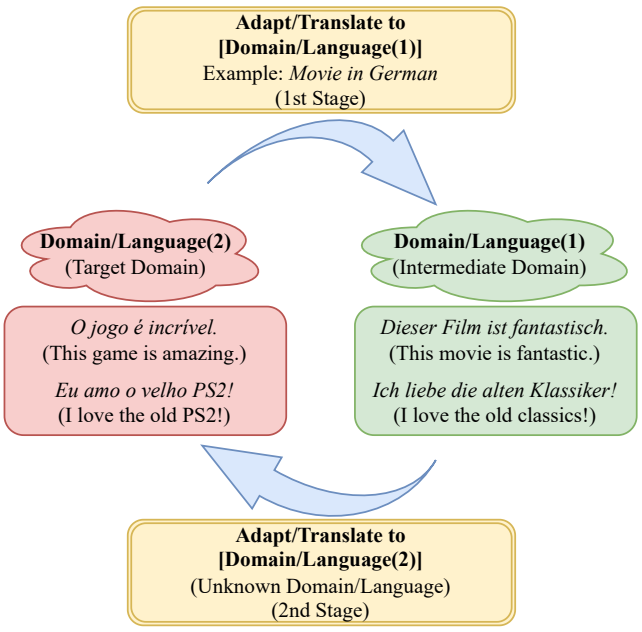


Figure 1. Proposed model for language and domain adaptation.

The second stage involves the training and adjustment of some of the prompt embeddings using the soft prompt technique. In this stage, the model takes the sentence generated in the first stage and adapts and translates it back to its original form using almost the same prompt. However, the embeddings associated with the specific domain and language tokens in the prompt are adjusted to accurately reproduce the original sentence during the backpropagation. In other words, when given the prompt "*Adapt and translate the above sentence into the domain and language: [Domain/Language(2)]*", the embeddings for [Domain/Language(2)] are adjusted to better get the original value.

Mathematically, in the second stage, we adopt $f(w, x') = g(w, x')$, where g is a frozen LLM model whose output depends on the input sentence transformed in the first stage, x', and the vectors w representing the domain and language, which are learned, to form the prompt. The loss function $\mathcal{L}$ follows the autoregressive form:

$$\mathcal{L} = -\frac{1}{T}\sum_{t=1}^{T} \log P(y_t \mid y_{<t}, x), \tag{1}$$

where T is the target sentence length and P represents the probability of the next token.

It is important to emphasize that the entire training process occurs independently of the source domain data. Therefore, the tokens encoding the target language and domain (acquired during training) must generalize to data from unknown source domains. Furthermore, as demonstrated later, these tokens can be used in other prompts.

Moreover, the selection of the intermediate domain plays a crucial role in the overall performance of the methodology. When the intermediate domain differs from the final domain in topic and genre, the model can learn to recognize and adapt to these differences effectively. However, if the intermediate domain is entirely unrelated to the target domain, it may

lead to the generation of highly divergent sentences, potentially compromising the effectiveness of the second adaptation stage. For example, in an extreme case where the intermediate domain consists solely of integer numbers, adaptation from a textual source becomes infeasible, and the learning process is likely to be severely impaired.

6 Experiments

This section provides a detailed description of the experiments conducted, including the datasets, baseline model, evaluation metrics, and model configuration.

6.1 Datasets

For the target domain, we selected two Portuguese datasets. The first dataset contains sentences in discussions about console and computer games on *Reddit* from the *gamesEcultura* group (*Games in Portuguese*). It was created by combining the titles and sentences from the comments, resulting in 3491 sentences, and is made available along with the source code of this paper. The second dataset is a Financial Twitter dataset with sentences from Brazilian Financial Market News (*Financial in Portuguese*) [Zerbinati *et al*., 2024]. It contains 4048 tweets annotated for named entity recognition. The annotations were ignored in this paper. *Financial in Portuguese* is a more challenging dataset to learn than the *Games in Portuguese*. Despite its specific terminology, the latter is more straightforward. The former is more difficult due to the frequent use of links and abbreviations. Thus, we aim to present the different results across different types of domains.

For the source domain, the test datasets of English comments on restaurants (*Restaurant in English*) and laptops (*Laptop in English*) from SemEval [Pontiki *et al*., 2016], annotated for Aspect-Based Sentiment Analysis, and the SEntFiN dataset of Indian Financial Market News Tweets (*Financial in English*), written in English, annotated for sentiment and aspect category extraction [Sinha *et al*., 2022], were selected. These datasets contain 800, 2158, and 2151 sentences, respectively. The dataset annotations were almost entirely ignored, except for a preliminary test for domain adaptation on labeled data described in this paper. These English datasets are frequently used as benchmarks for domain adaptation tasks in Sentiment Analysis. In particular, the *Financial in English* is a more challenging domain, with many abbreviations. Thus, we aim to present results for different levels of difficulty.

6.2 Baseline

The sentences generated with the learned language and domain should improve performance compared to those generated by an LLM model with pre-determined language and domain. For instance, for the *Games in Portuguese* dataset, the sentences generated by the model with the trained domain should match more closely than those generated by an LLM model with a pre-fixed prompt designed to adapt sentences to that domain.

For these experiments, the chosen Baseline is an LLAMA model, specifically an 8-billion-parameter variant trained for instruction. A predefined prompt is used to adapt sentences to the *Games in Portuguese* and *Financial in Portuguese* domains, depending on the target domain.

6.3 Evaluation Metrics

We analyzed the adaptation from both topic and genre perspectives. For the topic analysis adaptation, this study presents the results for two quantitative metrics:

- *Maximum Mean Discrepancy (MMD) with the Gaussian kernel* – used to measure the proximity between two distributions and commonly employed in domain adaptation studies [Zhang *et al*., 2023b; Yu *et al*., 2023]. It was used to determine how close the distribution of the generated sentences is to the target distribution. The *paraphrase-multilingual-MiniLM-L12-v2* model, a Sentence BERT [Reimers and Gurevych, 2019] trained on a paraphrase multilingual dataset, was used to generate a vector representation for each sentence. The paraphrase task compels the model to generate similar embeddings for different phrases with similar meanings. This model has presented good performance in Semantic Textual Similarity (STS) tasks [Muennighoff *et al*., 2022]. The configuration that achieves the lowest MMD is sought.
- *Standard Deviation (SD)* – the model could be generating only a single sentence, ignoring the original sentence. By measuring the SD, we aim to demonstrate that the generated sentences differ among them, indicating that the original sentences influence the generated sentences.

Note that these metrics measure two requirements of the formulated problem: the proximity between the generated and target distributions through MMD and whether the generated sentences maintain a relationship with the original sentence through SD in terms of meaning.

Saunders [2022] proposes to identify a text topic based on vocabulary distribution. Thus, we evaluate two additional metrics: a Jaccard similarity of the vocabulary (Vocab Jacc) of all sentences between the adapted and the target domain, and a Jensen-Shannon distance of the bi-gram distribution (2-gram JSD) to capture jargon. In this case, a better approximation has a high Jaccard similarity and a low Jensen-Shannon distance.

For the genre adaptation analysis, we consider the frequency distribution of syntactic categories. To quantify differences between source and target domains, we compute the Jensen-Shannon distance (JSD) for these distributions. Specifically, we measure the JSD for part-of-speech tags (POS JSD) and for syntactic dependency relations (Dep JSD), both extracted using the SpaCy[2] library. Although POS tags represent morphosyntactic elements, their distribution in a dataset can influence the overall syntactic structure of sentences, as they are closely related to the grammatical functions that words fulfill within syntactic constructions. Additionally, bi-gram statistics are included to complement the analysis of writing style, following Ríos-Toledo *et al*. [2022].

It is important to emphasize that metrics such as BLEU [Papineni *et al*., 2002], METEOR [Banerjee and Lavie, 2005],

[2] `https://spacy.io/`

and BertScore [Zhang *et al*., 2019], which are commonly used to evaluate translation models, are unsuitable in this context. Unlike translation tasks, where direct text pairs are available for comparison, obtaining such pairs for adaptation evaluation is not feasible. For example, consider adapting the restaurant-related sentence "*I love pudding*" to the *Movies in Portuguese* domain. Valid adaptations might include "*I love Star Wars*", "*I enjoy horror movies*", or "*I really like movies about food*". In this case, conventional translation metrics would fail to evaluate meaningfully, as they rely on strict lexical overlap rather than assessing semantic adequacy to the new topic.

6.4 Experimental Setup

As described in **Section 5**, the solution consists of two stages. In the first stage, sentences from the target domain are converted into any pre-established language and domain. The solution's performance depends on the languages and domains selected in this first stage. Therefore, several languages and domains were considered for this stage according to the target datasets:

- ***Games in Portuguese*** - (i) *Sport in English* (sp_EN); (ii) *Sport in German* (sp_DE); (iii) *Games in Portuguese* (ga_PT); and (iv) random conversations among *Sport in English*, *Celebrity in German*, and *Movie in Russian* (Several domains).
- ***Financial in Portuguese*** - (i) *Sport in English* (sp_EN); (ii) *Sport in German* (sp_DE); (iii) *Financial in Portuguese* (fi_PT); and (4) random conversations among *Sport in English*, *Celebrity in German*, and *Movie in Russian* (Several domains).

We aimed to create scenarios in which we varied the language, the domain, or both. The English language is closer to the source datasets but further from the target. The Portuguese language is closer to the target datasets but further from the source datasets. German and Russian are distant from both. Similarly, we tested the original domains (*Games* or *Financial*) as well as other domains such as *Sport*, *Celebrity*, and *Movie*.

Moreover, in the second stage, the initial embedding values corresponding to the target domain representation to be learned can be selected in various ways. We sought to create scenarios where we start with the target domain and language closer to or further from the desired one. The following options were tested:

- ***Games in Portuguese*** - (i) *Portuguese* (PT); (ii) *Games* (ga); (iii) *Games in Portuguese* (ga_PT); and (iv) *Movie* (mv).
- ***Financial in Portuguese*** - (i) *Portuguese* (PT); (ii) *Financial* (fi); (3) *Financial in Portuguese* (fi_PT); and (iv) *Movie* (mv).

The LLAMA 3 model with 8 billion parameters was used, with the instruction variant using a quantized version during training [Touvron *et al*., 2023]. The model was trained with the AdamW optimizer with an initial learning rate of 10^{-4} and a constant decay for five epochs. The choice of five epochs was made after observing that the loss function decline remained minimal after the first epoch (see **Appendix 10.2**). For the MMD metric, the gamma parameter was based on the median. The training batch size used was 18.

7 Results

The results are presented separately for the two target datasets: *Games in Portuguese* and *Financial in Portuguese*. For each dataset, two tables with all quantitative metrics are provided: (i) a comparison between the adapted dataset and the target dataset, and (ii) a comparison between a direct translation of the original datasets (Restaurant, Laptop, and Financial) into Portuguese using the large LLM *DeepSeek V3*[3], and the adapted dataset. Accordingly, the first table for each dataset (**Tables** 2 and 4) shows the similarity between the adapted and the target dataset, while the second table (**Tables** 3 and 5) demonstrates the similarity to the original datasets. An effective adaptation is expected to approximate the target dataset (first table) while diverging from the source datasets (second table). For didactic purposes, the results were averaged across different initializations of the final domain. Detailed results are provided in **Appendix 10.3**.

Table 2 illustrates the quantitative results for adapting the *Restaurant*, *Laptop*, and *Financial* datasets *in English* to *Games in Portuguese*. The table contains the result of our Baseline, i.e., the LLAMA model with a fixed prompt. The remaining rows present the results obtained for the intermediate domain (*Intermediate Domain*), averaged across different initializations of the target domain. **Table 3** illustrates the metrics between the original dataset translated to Portuguese and the same dataset adapted to the *Games in Portuguese* dataset.

The proposed model demonstrated superior proximity to the target distribution for the MMD metric under many configurations in *Games in Portuguese* adaptation, indicating a better proximity in meaning distribution. However, the model presented a poor performance for all metrics for the models with the intermediate domain *Games in Portuguese*. In this case, we believe this happens because, when using an intermediate domain close to the target domain (both are games in Portuguese), the model learns a neutral domain, i.e., it keeps the sentences unchanged in many cases. Thus, the model with the trained language and domain fails when applied to the *Restaurant*, *Laptop*, and *Financial in English* datasets.

An important exception was observed in the adaptation from the *Laptop* domain to the *Games* domain, which yielded lower performance across nearly all combinations of intermediate and final domains. One possible explanation is the substantial domain similarity, which could introduce significant challenges in the process. This similarity may result in frequent errors and, in some cases, cause the model to implicitly avoid adaptation. Supporting this hypothesis, the syntactic distance metrics—*POS JSD* and *Dep JSD*—for the adapted dataset were closer to those of the original *Laptop* dataset translated into Portuguese (Table 3) rather than to those of the Baseline, which contrasts with the behavior observed in the adaptation of other datasets.

The high *Vocab Jacc* values and low *2-gram JSD* scores obtained in the adaptation to the *Games in Portuguese* domain

[3] `https://www.deepseek.com/en`

Table 2. Adaptation to *Games in Portuguese*

Source Dataset	Itermediate Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
Restaurant	LLAMA 8B	0.1491	0.7303	0.1066	0.2275	0.0354	0.8100
Restaurant	Serveral domains	0.1466	0.7199	0.1062	0.2249	0.0302	0.8759
Restaurant	Sport in English	0.1496	**0.7185**	0.1017	0.2211	**0.0216**	0.8678
Restaurant	Sport in German	**0.1517**	0.7213	**0.0976**	**0.2150**	0.0262	0.8789
Restaurant	Games in Portuguese	0.1089	0.7567	0.1696	0.2846	0.0724	**0.8792**
Laptop	LLAMA 8B	**0.1129**	0.7381	0.0963	0.2284	0.0299	0.8136
Laptop	Serveral domains	0.1084	0.7357	0.0879	0.2122	0.0381	0.8689
Laptop	Sport in English	0.1077	**0.7340**	0.0897	0.2165	**0.0294**	0.8637
Laptop	Sport in German	0.1121	0.7364	**0.0805**	**0.2018**	0.0378	0.8702
Laptop	Games in Portuguese	0.0893	0.7637	0.1681	0.2859	0.0573	**0.8742**
Financial	LLAMA 8B	0.0982	0.7913	0.2910	0.3925	0.0685	0.8166
Financial	Serveral domains	0.1035	**0.7625**	**0.1798**	**0.2917**	0.0552	0.9111
Financial	Sport in English	0.1056	0.7671	0.2184	0.3211	0.0521	0.9058
Financial	Sport in German	**0.1100**	0.7664	0.2074	0.3138	**0.0474**	0.9066
Financial	Games in Portuguese	0.0557	0.8106	0.3630	0.4665	0.1079	**0.9227**

Note: Bold values indicate the best result per source domain. Metrics cover lexical overlap (Vocab Jacc), stylistic shift (2-gram), syntactic divergence (POS JSD, Dep JSD), semantic distance (MMD), and variation in generated sentences (STD).

Table 3. Pure Portuguese Translation (Source Dataset) vs. Adaptation to *Games in Portuguese* Domain

Source Dataset	Itermediate Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	Src STD
Restaurant	LLAMA 8B	**0.2465**	**0.6895**	0.0532	0.1338	**0.1204**	**0.8526**
Restaurant	Serveral domains	0.3032	0.6623	0.0720	0.1532	0.0387	**0.8526**
Restaurant	Sport in English	0.2712	0.6813	0.0726	0.1629	0.0572	**0.8526**
Restaurant	Sport in German	0.3163	0.6523	0.0572	0.1401	0.0380	**0.8526**
Restaurant	Games in Portuguese	0.2670	0.6838	**0.1847**	**0.2732**	0.0044	**0.8526**
Laptop	LLAMA 8B	0.3267	0.6610	0.0778	0.1831	**0.0861**	**0.8646**
Laptop	Serveral domains	0.3786	0.6254	0.0611	0.1670	0.0137	**0.8646**
Laptop	Sport in English	0.3391	0.6533	0.0652	0.1824	0.0250	**0.8646**
Laptop	Sport in German	0.3988	0.6138	0.0496	0.1593	0.0120	**0.8646**
Laptop	Games in Portuguese	**0.2827**	**0.6715**	**0.1636**	**0.2775**	0.0021	**0.8646**
Financial	LLAMA 8B	**0.2481**	0.7508	0.1367	0.2279	**0.1070**	**0.9065**
Financial	Serveral domains	0.2840	0.7472	**0.2255**	0.3139	0.0287	**0.9065**
Financial	Sport in English	0.2521	**0.7518**	0.1888	0.2877	0.0334	**0.9065**
Financial	Sport in German	0.2721	0.7413	0.1909	0.2744	0.0344	**0.9065**
Financial	Games in Portuguese	0.2953	0.7451	0.2253	**0.3632**	0.0032	**0.9065**

Note: Bold values indicate the best result per source domain. Metrics cover lexical overlap (Vocab Jacc), stylistic shift (2-gram), syntactic divergence (POS JSD, Dep JSD), semantic distance (MMD), and variation in pure translated source sentences (STD).

suggest that the generated sentences moved away from the source domain while becoming more aligned with the target domain (**Table 2**), outperforming the Baseline (LLAMA 8B) in this aspect. This indicates a more effective domain adaptation. There are two possible explanations for this behavior: (i) the proposed method may be memorizing frequent words from the target domain, or (ii) the method may be successfully generating sentences with vocabulary more characteristic of the target domain. In both scenarios, the high standard deviation (STD) of the generated outputs (**Table 2**) suggests considerable variation in meaning, further supporting the occurrence of meaningful adaptation. All models presented higher values of SSD than the baseline model, indicating a considerable variation among the generated sentences. It is important to note that this metric should be analyzed with the MMD, as the model might have generated disconnected sentences in the case of a high MMD and standard deviation.

Regarding the syntax, a similar behavior was observed. The proposed method presented a closer distribution for the *POS-tag* (POS JSD) and *syntax dependency tree* (Dep JSD) (**Table 2**) and diverging further from the source domain (**Table 3**) compared to the Baseline. **Table 3** compares the results with the pure Portuguese translation; thus, we can conclude that the divergence between the original dataset is not only due to a different syntax in the Portuguese language.

The approximation for the adaptation to the *Financial in Portuguese* domain is illustrated in **Table 4**. Similar to the adaptations for the "Games in Portuguese" dataset, the model has the poorest performance with its homonymous "Finance in Portuguese" as intermediate domain. But, even in this case, it presented a better approximation in *MMD*, *POS JSD*, and *Dep JSD* than the Baseline *LLAMA 8B*. This is possible because the intermediate adaptation to *Financial in Portuguese* during the training does more than only copying the original sentence; in other words, it changes the original sentence, indicating that the LLM has a different understanding of the original dataset.

The *Vocab Jacc* metric in **Table 4** suggests that the proposed model was less effective than the Baseline in performing lexical transformation, which may have hindered topic adaptation. This interpretation is supported by **Table 5**, where the lexical distances (*Vocab Jacc*) relative to the original dataset remain similar across both approaches applied to the Restaurant and Laptop source datasets. In other words, although both approaches modified the lexicon, the Baseline achieved more effective lexical adaptation. However, the MMD points in the opposite direction, showing a more proximal distribution. This indicates that, despite the difference in the lexical distribution, the proposed method generated sentences that have a more similar distribution of semantic meaning.

As observed in the *Games in Portuguese* dataset, the results illustrated for the syntactic analysis in *POS JSD* and *Dep JSD* in **Table 4** show a better approximation for the proposed model in most of the cases. Finally, despite the lexical having a poor performance, the *2-gram JSD* presents a better distribution approximation, which can suggest a good style approximation.

Regarding the *Initial Final Domain*, the adaptation presented slightly better results in many cases when the initial value for the final domain was closer to the specification. This was mainly noticed for the *Games in Portuguese* domain, showing the importance of choosing a closer start. Additional details can be found in **Appendix 10.3**.

For a more detailed understanding of the model adaptation, we build a graph comparing the top 10 nouns most cited in the targets domains (Original), *Games in Portuguese* (**Figure 2**) and *Financial in Portuguese* (**Figure 3**), with the top 10 nouns most cited in the adaptations using the *LLAMA 8B* and the *Learned Domain*. For the latter, we excluded the learned domains that used *financial in Portuguese* and *games in Portuguese* as intermediate domains. For this process, we use the *Spacy* library, which considers some special symbols as nouns (example: "@").

For the *Games in Portuguese* domain, words such as "jogo" (game), "jogos" (games), and "jogadores" (players) were cited approximately twice as often in sentences adapted by the *LLAMA 8B* model compared to those in the original dataset. We hypothesize that the baseline model (LLAMA 8B), lacking specific knowledge of the target domain, tends to generalize the vocabulary toward a broader "games" topic. Beyond, the frequent presence of terms like "estratégia" (strategy) and "futebol" (football) suggests that, in some cases, the baseline interprets "games in Portuguese" as referring to sports rather than video games. In contrast, the *Learned Domain* produces adapted sentences with a word distribution more aligned with the source material.

A similar pattern was observed in the *Financial in Portuguese* domain. Terms such as "investimento" (investment), "gestão" (management), and "mercado" (market) appeared more often in the sentences adapted by the *LLAMA 8B*, whereas original and *Learned Domain* sentences include more specific terms like "petr4" and "vale5". Once again, the *Learned Domain* demonstrates a more faithful lexical distribution. However, a noteworthy issue arises with the absence of elements such as the link "`http://t.co/kgtlyibf7`" and the tag "@live_trade"-frequent in the original data but missing from the adapted outputs. During training, such links and tags are removed in the intermediate domain stage, making it impossible to fully recover them in the final adaptation. This limitation may lead to inconsistencies in the learning process and could explain the anomalous presence of isolated "@" symbols in many sentences adapted using the *Learned Domain*.

8 Discussion

8.1 Qualitative Analysis

We qualitatively analyze the proposed method from three perspectives: (i) the generation of coherent sentences in Portuguese; (ii) the approximation to the target domain from multiple angles; and (iii) sentence transformations related to adaptation and translation. To support this analysis, we employ GPT-4.1, a large language model with advanced linguistic capabilities. GPT-4.1 is an improved version of GPT-4, offering broader coverage and enhanced reasoning abilities [OpenAI, 2024]. While syntactic evaluations in previous studies were conducted on earlier models such as GPT-3.5

Table 4. Adaptation to *Financial in Portuguese*

Source Dataset	Itermediate Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	Src STD
Restaurant	LLAMA 8B	**0.0902**	0.7919	0.2619	0.4075	0.1624	0.8343
Restaurant	Serveral domains	0.0852	0.7857	0.1525	0.3275	0.0626	0.8399
Restaurant	Sport in English	0.0828	0.7833	**0.1374**	**0.3160**	0.0659	0.8537
Restaurant	Sport in German	0.0849	**0.7815**	0.1644	0.3253	**0.0583**	**0.8564**
Restaurant	Financial in Portuguese	0.0595	0.8042	0.1672	0.3590	0.1250	0.8358
Laptop	LLAMA 8B	**0.0677**	0.7973	0.2757	0.4287	0.1697	0.8372
Laptop	Serveral domains	0.0617	0.7942	0.1499	0.3243	0.0591	0.8429
Laptop	Sport in English	0.0565	0.7941	**0.1357**	**0.3189**	0.0637	**0.8642**
Laptop	Sport in German	0.0620	**0.7909**	0.1776	0.3379	**0.0590**	0.8619
Laptop	Financial in Portuguese	0.0451	0.8088	0.1685	0.3583	0.1137	0.8510
Financial	LLAMA 8B	**0.1003**	0.7884	0.2100	0.3492	0.1187	0.8617
Financial	Serveral domains	0.0791	0.7940	0.2026	0.3493	0.0542	0.8208
Financial	Sport in English	0.0831	0.7868	0.1696	**0.2937**	0.0474	0.8640
Financial	Sport in German	0.0865	**0.7814**	**0.1653**	0.3008	**0.0379**	0.8232
Financial	Financial in Portuguese	0.0651	0.8024	0.1912	0.3255	0.0488	**0.8651**

Note: Bold values indicate the best result per source domain. Metrics cover lexical overlap (Vocab Jacc), stylistic shift (2-gram), syntactic divergence (POS JSD, Dep JSD), semantic distance (MMD), and variation in generated sentences (STD).

Table 5. Pure Portuguese Translation (Source Dataset) vs. Adaptation to *Financial in Portuguese* Domain

Source Dataset	Itermediate Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
Restaurant	LLAMA 8B	0.2173	0.7116	0.0749	0.1570	**0.1567**	**0.8526**
Restaurant	Serveral domains	0.2201	0.7193	0.1874	0.3440	0.1146	**0.8526**
Restaurant	Sport in English	0.2163	0.7222	0.1812	0.3171	0.0943	**0.8526**
Restaurant	Sport in German	0.2256	0.7214	0.1737	0.3184	0.0881	**0.8526**
Restaurant	Financial in Portuguese	**0.1818**	**0.7419**	**0.2652**	**0.4151**	0.0824	**0.8526**
Laptop	LLAMA 8B	0.2546	0.6997	0.1292	0.2345	**0.1204**	**0.8646**
Laptop	Serveral domains	0.2658	0.7044	0.1609	0.3376	0.1085	**0.8646**
Laptop	Sport in English	0.2503	0.7100	0.1636	0.3233	0.0762	**0.8646**
Laptop	Sport in German	0.2678	0.7117	0.1677	0.3372	0.0788	**0.8646**
Laptop	Financial in Portuguese	**0.1949**	**0.7340**	**0.2479**	**0.4114**	0.0808	**0.8646**
Financial	LLAMA 8B	0.2787	0.7165	0.1047	0.1937	0.0372	**0.9065**
Financial	Serveral domains	0.2337	0.7625	**0.2565**	**0.4394**	**0.0906**	**0.9065**
Financial	Sport in English	0.2549	0.7499	0.1871	0.3377	0.0376	**0.9065**
Financial	Sport in German	**0.2264**	**0.7626**	0.2190	0.3784	0.0705	**0.9065**
Financial	Financial in Portuguese	0.2281	0.7615	0.2402	0.4082	0.0396	**0.9065**

Note: Bold values indicate the best result per source domain. Metrics cover lexical overlap (Vocab Jacc), stylistic shift (2-gram), syntactic divergence (POS JSD, Dep JSD), semantic distance (MMD), and variation in pure translated source sentences (STD).

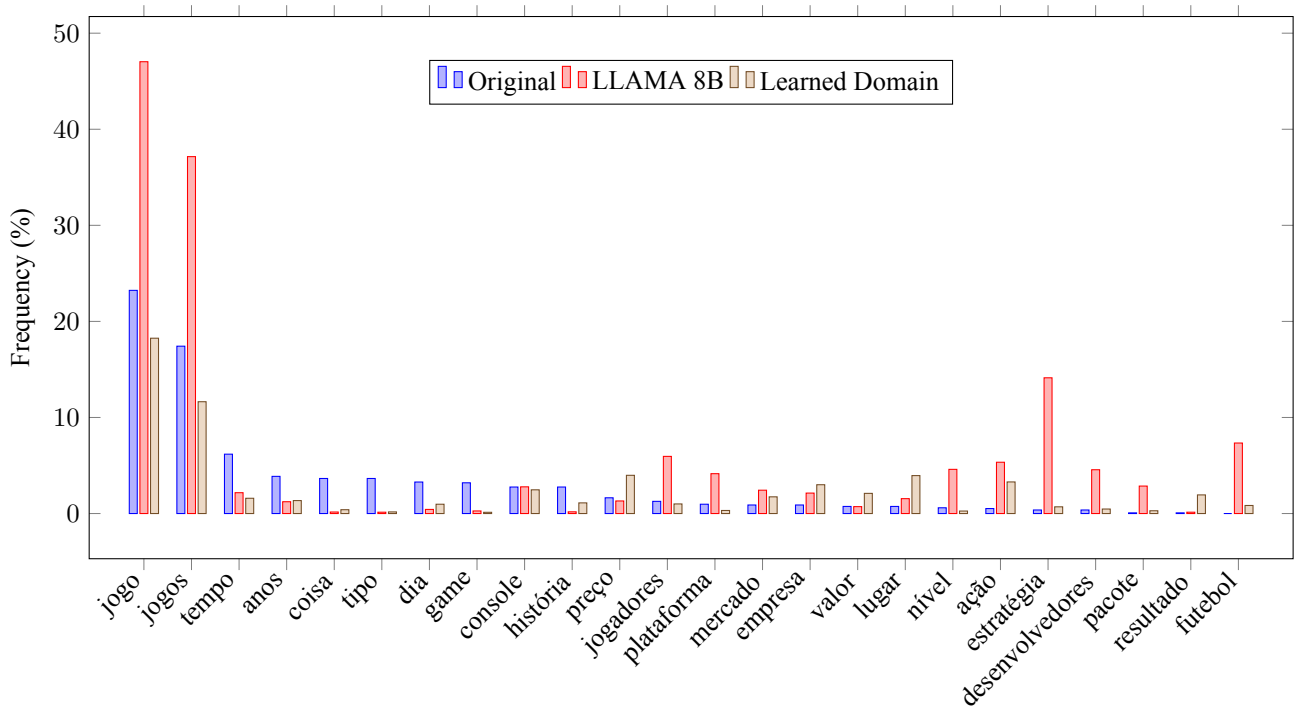


Figure 2. The Top 10 most cited words in the original and adapted sentences - Games in Portuguese

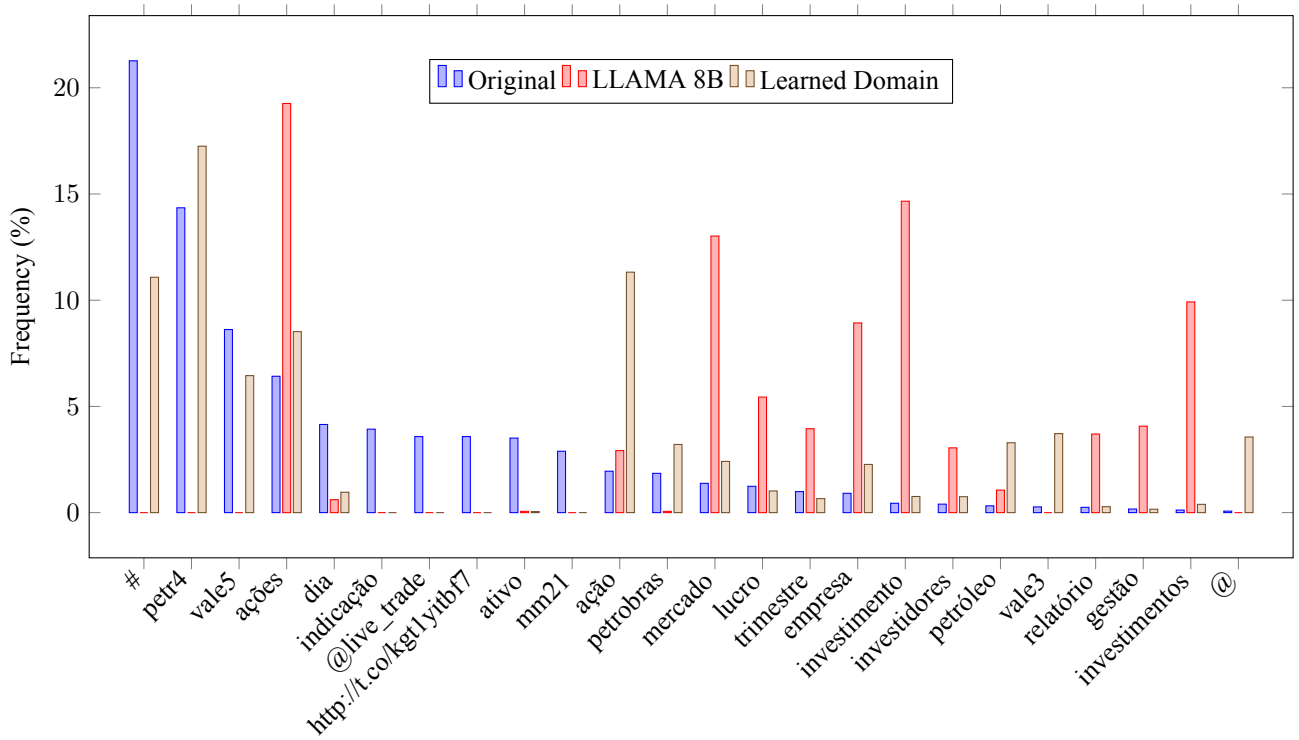


Figure 3. The Top 10 most cited words in the original and adapted sentences - Financial in Portuguese

and GPT-4 [Zhou *et al*., 2023], GPT-4.1 builds upon these foundations, providing more robust language understanding.

While we recognize the limitations inherent in automated evaluations, this approach enables a detailed and consistent review across thousands of sentence pairs—a task that would be prohibitively time-consuming and less consistent if performed manually. To ensure reliability, the initial examples were validated by a human annotator. This strategy, commonly referred to in the scientific literature as "*LLM as a judge*", has demonstrated promising results when leveraging very large language models such as GPT-4.1 [Zheng *et al*., 2023; Gu *et al*., 2024]. We provide both the original and adapted datasets to support and enhance manual curation by human experts. The data are available from the authors upon reasonable request, in accordance with relevant ethical considerations.

All models were trained on a training subset and evaluated on a held-out test subset. Due to budget constraints, we limited our evaluation with the LLM to two representative combinations of source, intermediate, and target domains for the proposed method:

- Adaptation from *Laptop in English* to *Games in Portuguese*, using *Sport in English* and *Games* as the intermediate and initial target domains, respectively. This scenario was selected due to its poor results in *POS JSD* and *Vocab Jacc* metrics (**Table 2**). The test dataset has 800 sentences.
- Adaptation from *Financial in English* to *Financial in Portuguese*, using *Sport in German* and *Financial in Portuguese* as the intermediate and initial target domains, respectively. This case was chosen because it yielded the best results for almost all metrics (**Table 4**). The test dataset has 2158 sentences.

From the first perspective, we examine and compare the coherence and language of the adapted sentences produced by the proposed method (Learned Domain) and the Baseline (LLAMA 8B). The exact prompt used for this evaluation is provided in **Appendix 10.4**. **Tables** 6 and 7 present the results for both models in the target domains *Games in Portuguese* and *Financial in Portuguese*, respectively.

Table 6. Coherence and Language of the adapted sentences - Games in Portuguese

Model	Coherence	Port.	Both
LLAMA 8B	**96.38%**	**99.75%**	**96.12%**
Learned Domain	83.62%	99.12%	82.88%

Table 7. Coherence and Language of the adapted sentences - Financial in Portuguese

Model	Coherence	Port.	Both
LLAMA 8B	**98.33%**	**97.77%**	**96.14%**
Learned Domain	77.78%	81.59%	69.27%

The proposed model displays notable coherence issues and underperforms compared to the baseline when adapting and translating sentences into Portuguese. We hypothesize that this limitation arises from a specific aspect of the training process: during the transition to the intermediate domain, the model tends to lose essential information—such as named entities, abbreviations, tags, and links—and is subsequently penalized when mapping back to the target domain. In an attempt to compensate for this loss, the model often generates new content, which is frequently spurious and introduces noise into the learned domain representation. This issue is particularly pronounced in the *Financial in Portuguese* domain, which contains a higher density of tags, links, and abbreviations, making it more susceptible to degradation than the *Games in Portuguese* domain.

The second perspective to be analyzed is the approximation of the adapted sentences to the destination domain. For this, we use the LLM to indicate which phrase adapted has more similar domain attributes to the 20 sentences randomly sampled from the destiny domain. We observed that comparing the sentences adapted by the two approaches (Learned Domain and LLAMA 8B), instead of testing them individually, can present less inconsistent answers. The exact prompt is in **Appendix 10.5**. **Tables** 8 and 9 present the comparison between the proposed method (Learned Domain) and the Baseline (LLAMA 8B). For this analysis, we removed the sentences generated with meaning and translation problems.

Table 8. Comparison between the Learned Domain (L.D) and the Baseline (LLAMA) - Games in Portuguese

Item	L.D.(%)	LLAMA (%)	Tie (%)
Topic	**91.04**	2.78	6.18
Style	24.27	**70.79**	4.95
Syntactic Structure	**76.97**	15.30	7.73
Named entities	**20.25**	6.80	72.95
Sentiment	**31.07**	4.48	64.45
General meaning	**48.84**	14.53	36.63
Function	**64.61**	6.49	28.90
Register	**47.45**	3.86	48.69

Note: Bold values indicate the best result per method, ignoring ties.

Table 9. Comparison between the Learned Domain (L.D) and the Baseline (LLAMA) - Financial in Portuguese

Item	L.D.(%)	LLAMA (%)	Tie (%)
Topic	**99.40**	0.40	0.20
Style	**49.13**	41.74	9.13
Syntactic Structure	**90.54**	3.49	5.97
Named entities	**90.40**	2.68	6.91
Sentiment	**3.69**	0.67	95.64
General meaning	30.40	**32.62**	36.98
Function	**44.23**	8.66	47.11
Register	**82.28**	5.84	11.88

Note: Bold values indicate the best result per method, ignoring ties.

The proposed model was selected by the LLM as a better approximation for nearly all domain attributes, except for the style attribute in the *Games in Portuguese* domain and the general meaning attribute in the *Financial in Portuguese* domain. In the first case, contrary to what was proposed by Ríos-Toledo *et al*. [2022], the quantitative results—particularly the bi-gram metric for style analysis shown in **Tables** 2 and 4—diverge from the qualitative evaluation. This suggests that the bi-gram metric may be insufficient to assess stylistic

features in the proposed method for certain domains. In the second case, the source domain (*Financial in English*) and the target domain (*Financial in Portuguese*) are more closely aligned in terms of meaning, making the adaptation process easier. As a result, the models perform similarly, with the LLAMA model achieving a slight advantage.

The *topic* metric emerges as the most discrepant among all evaluated attributes. In the *Games in Portuguese* adaptation, it reaches 91.04% in favor of the proposed method, compared to just 2.78% for the baseline, and achieves 99.40% in the *Financial in Portuguese* domain. When considered alongside the *Vocab Jac* metric—presented for the Laptop group in **Table 2** and for the Financial group in **Table 4**—where the baseline model shares a larger number of words, we observed that the Learned Domain adopts a smaller, more domain-aligned vocabulary. In contrast, the Baseline model employs a broader vocabulary that, while more lexically diverse, often strays from the linguistic patterns characteristic of the target domain. This interpretation is further supported by the *Named Entities* metric, which also favors the Learned Domain.

The *General meaning* metric shows a smaller discrepancy compared to *Topic*, which reinforces the hypothesis that the Baseline model employs a broader vocabulary. This is evidenced by the greater number of ties with the proposed method, indicating a possibility that, although different words are used, the conveyed meaning remains similar. One possible explanation for the observed difference in the *Games in Portuguese* domain is the ambiguity of the term "games", which may refer either to digital games or to sports. The latter represents a domain quite distinct from the target, potentially contributing to the Baseline model's poorer performance in this context.

The values for *Syntactic structure* and *Register* indicate that the Learned Domain was able to learn some structure of the sentences, forcing them to be restructured when adapting. Sentiment was the value with the least difference between the models, which is reasonable, since it should be preserved. Conversely to the expected, *Function* presents a significant difference between the Learned Domain and the LLAMA, which makes us believe that the LLM specialist was very rigorous in the evaluation to suggest that the Learned Domain has a better performance.

For the third perspective, we analyze how the sentences were changed to meet the target domain. To evaluate this, we consult an LLM twice, once for the Baseline (LLAMA 8B) and once for the proposed method (Learned Domain), asking for the domain attributes changed in the adaptation. **Tables** 10 and 11 illustrate the percentage of sentences changed for each attribute for both methods.

Some attributes, such as *Sentiment* and *Function*, exhibited minimal variation, which was expected. However, the value associated with the *Function* attribute in this third perspective, **Tables** 10 and 11, appears to contradict the results presented in the second perspective, **Tables** 8 and 9. From a second perspective, the large LLM presents a preference in many cases between the Baseline and the proposed method; therefore, the function must be varied. We hypothesize that, in that case, the LLM performed a more fine-grained analysis to determine that the *Learned Domain* is closer in terms of *Function*. As a result, the change in the function was subtle but present, being ignored by the large LLM in this third perspective.

Table 10. Similarity between the source domain and the adapted sentences for the Learned Domain and the Baseline (LLAMA 8B) - Games in Portuguese

Item	Learned Domain	LLAMA 8B
Topic	33.94%	1.96%
Style	23.38%	34.54%
Syntactic Structure	39.67%	52.49%
Named entities	77.83%	77.83%
Sentiment	92.01%	95.17%
General meaning	40.12%	12.82%
Function	92.01%	96.08%
Register	49.32%	44.34%

Table 11. Similarity between the source domain and the adapted sentences for the Learned Domain and the Baseline (LLAMA 8B) - Financial

Item	Learned Domain	LLAMA 8B
Topic	27.72%	16.44%
Style	7.18%	9.40%
Syntactic Structure	25.64%	38.52%
Named entities	19.93%	17.52%
Sentiment	96.51%	95.30%
General meaning	36.17%	22.89%
Function	99.13%	99.33%
Register	58.52%	77.32%

The differences observed between **Table 10** and **Table 11** suggest that the adaptation process to the target domain undergoes more substantial changes when moving from *Financial in English* to *Financial in Portuguese* than from *Laptop in English* to *Games in Portuguese* for some attributes, such as *Named entities* and *Topic*. Although this may initially seem contradictory, the *Games in Portuguese* domain contains numerous sentences related to computer games, which often mention Laptops—creating a natural overlap with the source domain. In contrast, while both the *Financial in English* and *Financial in Portuguese* domains address financial topics, the former includes many sentences from the Indian market with minimal use of tags or links, whereas the latter is composed largely of posts from X (formerly Twitter) focused on the Brazilian market, frequently containing links, hashtags, and abbreviations, what could affect syntax or even style, once the sentences are individually compared.

8.2 Study Case

For this case study, we present a comparison between sentences adapted using a fixed-domain LLM (LLAMA 8B) and the proposed method (Learned Domain). We evaluated a case of one sentence adapted to *Games in Portuguese* by both methods and another sentence adapted to *Financial in Portuguese*. For these considerations, we present the analysis made by GPT 4.1, reviewed by a human. The prompt is in **Appendix 10.7**.

Table 12 illustrates the adaptation of the sentence "*UP Elections: Sugar stocks like Bajaj Hindusthan, Dhampur Sugar rise amid election results*", originally from the *Financial in English* domain, to the *Games in Portuguese* domain. The model used the combination of *Sport in German/Finan-*

cial in Portuguese to learn. The sentence generated by the model with the learned domain is more similar to that of an online discussion forum, which aligns with the target domain. On the other hand, the sentence generated by the model with a fixed domain confused video games with sports games. Additionally, it generated a sentence resembling a headline, which is more aligned with the source domain than the target one.

Table 12. Comparison of sentences generated by LLM for Games in Portuguese with fixed vs. learned domain

Model	Phrase
Original	UP Elections: Sugar stocks like Bajaj Hindusthan, Dhampur Sugar rise amid election results
LLAMA 8B	*Jogos de futebol: Stocks de jogadores como Neymar, Messi e Ronaldo sobem em valor após resultados da Copa do Mundo* (Football games: Player stocks like Neymar, Messi, and Ronaldo rise in value after World Cup results)
Learned Domain	*Eles estão subindo os preços dos jogos de PS5 e Xbox Series X* (They are raising the prices of PS5 and Xbox Series X games)

Table 13 illustrates the analysis made by GPT 4.1, reviewed by a human. Sentence A is the LLAMA 8B adaptation, and Sentence B is the Learned Domain adaptation. For this specific sentence, the proposed model (sentence B) generated sentences closest to the target for all domain attributes.

Table 13. Comparison of Adapted Sentences Against Target Domain - Games in Portuguese

Attribute	**Better**	**Reasoning (Short)**
Topic	B	About digital games/consoles, not sports
Style	B	Conversational, matches target informality
Syntactic Structure	B	Simple, direct structure like targets
Named entities	B	PS5/Xbox central in target, not footballers
Sentiment	B	Matches mild complaint/discontent in domain
General meaning	B	Game industry news, not sports markets
Function	B	Informative blurb, matches target function
Register	B	Informal, oral, community-match

Note: Sentence A is the LLAMA 8B, and sentence B is the Learned Domain.

Table 14 presents the adaptation of the sentence "*We arrived and were seated immediately, which made us both happy*" from the *Restaurant in English* domain to the *Financial in Portuguese* domain. The sentence generated by the adapted domain incorporated references to stocks and companies, a common feature in most sentences within the target domain. Additionally, the sentence generated by the fixed-domain model has a more commentary-like tone, similar to the source domain but distinct from both the target domain and the sentences generated with the learned domain.

Table 14. Comparison of sentences generated by LLM for Financial in Portuguese with fixed vs. learned domain

Model	Phrase
Original	We arrived and were seated immediately, which made us both happy
LLAMA 8B	*O investimento foi feito com sucesso e os resultados foram apresentados imediatamente, o que nos deixou ambos satisfeitos* (The investment was successfully made, and the results were presented immediately, which left both of us satisfied)
Learned Domain	*O mercado de ações da Petrobras (PETR4) e da Vale (VALE3), que são as maiores da Bolsa, também estão em alta* (The stock market of Petrobras (PETR4) and Vale (VALE3), which are the largest on the exchange, is also on the rise)

Table 15 illustrates the analysis made by GPT 4.1, reviewed by a human. For this sentence, the proposed model (sentence B) generated sentences closest to the target for many domain attributes, but not all.

Table 15. Comparison of Adapted Sentences Against Target Domain - Financial in Portuguese

Attribute	**Better**	**Reasoning (Short)**
Topic	B	Explicitly about stocks/entities in target domain
Style	B	Matches news/social/finance tweet style
Syntactic Structure	A	Shares "immediately... which made us happy" structure with some target cases
Named entities	B	Uses frequent target domain entities (stocks/companies)
Sentiment	A	Explicit subjective reaction (mutual satisfaction), matching rare but present target elements
General meaning	Tie	Both are only partially representative
Function	B	Reporting/informing dominates the target
Register	B	Correct register/lexicon for the financial tweet/news domain

Note: Sentence A is the LLAMA 8B, and sentence B is the Learned Domain.

8.3 Nonsensical Sentences

A critical aspect of our evaluation is determining whether the model generates semantically meaningful sentences. A first observation is that the fixed-domain model mostly generates sentences that make sense for both target domains, i.e., *Games in Portuguese* and *Financial in Portuguese*. On the other hand, the learned domain model maintains this ability for the *Games in Portuguese* domain but makes significant errors for the Financial domain in Portuguese.

We evaluated the first 50 sentences of the *Financial in Portuguese* adaptation with errors, generated using a *Learned Domain* obtained through the combination of *Sports in German* as an intermediate domain, *Financial in Portuguese* as the initial target domain, and *Financial in English* as the source dataset. The evaluation was carried out in two stages. First, we prompted *DeepSeek R1*, a large language model with reasoning capabilities, to perform an initial categorization. Subsequently, the same 50 sentences were re-categorized by a human annotator for comparison and validation purposes. **Table 16** illustrates the results.

Table 16. Percentage distribution of error types, considering a total of 50 occurrences

Category	Percentage (%)
Only tags/links/words repetitions	34%
Excessive repetition	28%
Excessive similar tags/links/words	12%
Incomplete translation	38%
Context inconsistency	32%
Formatting (URL)	36%
Other formatting issues	6%
Truncated text	12%
Logic errors / Redundancy	14%

Note: The sum is not 100% because more the one error can be present in a sentence.

The categories are:

- **Only tags/links/words repetitions**: only or almost only repetitions of words, tags, or links, and no other text. Example: “EDP - EDP - EDP - EDP”.
- **Excessive repetition**: many tags/links/words repetitions. Example: “O modelo de comissionamento de compra e venda tem valor, mas taxas muito altas, diz o CEO da BM&FBOVESPA, Pedro Baruselli. #BMFBovespa #BMFBovespa”.
- **Excessive similar tags/links/words**: a small variation in the tags/links/words, but with excessive repetition. Example: “#VALE3 #VALE #VALE5 #VALE6”.
- **Incomplete translation**: translation errors on terms and structures from other languages to Portuguese. Example: “em a” (in the) instead of “na” or “#VALE5 Q4 PAT at R$ 18,68 milhões” (#VALE5 Q4 PAT at R$18.68 million) instead of “#VALE5 LL do 4T em R$ 18,68 milhões”.
- **Context inconsistency**: experts, companies, and Indian indexes (e.g., SEBI, RBI, Ashwani Gujral) were not replaced with Brazilian equivalents. Example: “@MiteshThacker SBI com alvo de R$ 2560 : Mitesh Thacker”.
- **Formatting (URL)**: bad URL formatting. Example: “http://t.co/4Q7Q7Q7Q7Q7Q”. Important: We ignore this type of error in **Tables** 6 and 7, but we have decided to include it in this analysis because it provides an interesting overview of the adaptation.
- **Other formatting issues**: Numerical issues, incorrect spacing, punctuation, etc. Example: “1. 194 %”
- **Truncated text**: Abruptly cut-off sentences (incomplete words or lack of conclusion). Example: “Ações de a Petrobras (PETR4) e a Vale (VALE3) subiram em a Bolsa de São Paulo em a sessão de hoje. A Petrobras fechou em R$ 14, 45 e”
- **Logic errors / Redundancy**: Nonsensical sentences, contradictions, or incoherent repetitions. Examples include: “Petrobras assume comando de Petrobras” (Petrobras takes control of Petrobras), which is redundant, and the phrase “imposto de serviço” (service tax) in the sentence “Investidores de ações da Petrobras podem enfrentar surpresa de imposto de serviço” (Share investors in Petrobras may face a surprise service tax), referring to a tax that does not exist in Brazil.

The most prominent error was *Incomplete translation*, referring to incorrect translations of terms and sentence structures from English to Portuguese within this domain. A potential solution for this issue is to fix the target language explicitly in the prompt and focus the adaptation process on learning the remaining attributes since the language itself is often already well-defined.

Learning both language and domain for translation and adaptation depends on the LLM’s ability to convert sentences to an intermediate domain. In other words, if the LLM fails to perform a good intermediate adaptation, the learned language and domain will be significantly impacted. In the case of *Financial in Portuguese* domain, the sentences are short, and links and hashtags prevail. When the intermediate transformation removes these items, the final transformation needs to reinvent them, causing the learned domain to incorporate characteristics to generate tags and links. As a result, it is common to encounter errors of only tags/links/words, many repetitions, and trying to simulate the tags or even the devised link with many letter and number repetitions.

Context inconsistency and *Logic errors/Redudancy* can present a problem in this methodology. The model penalizes when the adaptation can not guess the named entities. Thus, the learned domain forces the model to guess some common named entities, creating *Context Inconsistency* and *Logic errors/Redundancy*. A future improvement in the methodology can be to remove the named entities from the training. Therefore, re-engineering the intermediate and final prompts and better data handling are required for improved results.

8.4 Visualization

Figure 4 and **Figure 5** display the UMAP projection of sentence vectors from the source and target domains (blue and red colors, respectively), considering the data without adaptation, adapted using a fixed language and domain, and using the learned language and domain. The *paraphrase-multilingual-MiniLM-L12-v2* Sentence BERT

model [Reimers and Gurevych, 2019] was used to generate each vector, which means that sentences with similar embeddings must have similar meanings. The first figure illustrates the domains of *Games in Portuguese* and *Financial in English*. The second figure illustrates the domains of *Financial in Portuguese* and *Restaurant in English*. **Appendix 10.1** contains additional visualizations.

In the first visualization (**Figure 4**), the learned domain representation results from combining *Sport in English* as the intermediate domain with *Games in Portuguese* as the initial embedding target. In the second visualization (**Figure 5**), the learned domain is obtained by combining *Sport in German* as the intermediate domain and *Financial in Portuguese* as the initial embedding. Other domain combinations yield similar graph structures.

In the first subplot of each figure, the sentences have not transformed, and their vectors are entirely separated. The second subplot shows the sentences after transformation using a fixed domain. While they come closer together, there is still a significant distinction. Finally, the third subplot illustrates the sentences after transformation by the learned domain, where the distributions of the generated sentences and the target domain overlap.

Contrary to the qualitative analysis findings for the adaptation from *Financial in English* to *Financial in Portuguese* (Table 11), the Baseline model (*Fixed Domain*) presents poorer similarity in general meaning to the target domain when adapting from the *Restaurant in English* domain, compared to the *Learned Domain*. A smaller difference is observed in the transformation from *Financial in English* to *Financial in Portuguese*, as shown in Appendix 10.1.

8.5 Applications to Other Tasks

The learned language and domain vectors can be used for other tasks. As a preliminary exercise, they were applied in experiments to translate and adapt labeled datasets. For this, the prompt was rewritten, and the learned vectors representing the domain and language were appended. A simplified version of the prompt is: "*Above is a sentence with its corresponding labels for the Aspect-Based Sentiment Analysis (ABSA) task. Translate and adapt it to the language/domain: [Domain/Language(2)]*", where [Domain/Language(2)] represents the learned vectors. The qualitative analysis showed promising initial results. **Table 17** illustrates an example of a labeled laptop sentence converted to the *Games in Portuguese* domain.

To analyze the sentence adaptation, we employed GPT-4.1 as a domain specialist to identify which elements were preserved and which were modified in the adapted sentence. The prompt used is in **Appendix 10.9**. The results of this analysis are presented in **Table 18**. The analysis indicates that the adaptation primarily involved changes in topic — such as the inclusion of different named entities — as well as modifications in style and register.

Another task is to adapt a text for a marketing campaign proposal. In this case, imagine someone who intends to post a message about a product in the *Games in Portuguese* forum, preserving the communication style. For this exercise, we generate the prompt: "*Marketing campaign adaptation activity. Above is a text about a new product. We want to talk about this in a forum. Adapt the text above to generate a new text for the domain/language: [Domain/Language(2)]. VERY IMPORTANT: Don´t create any new information to adapt. Only adapt.*". **Table 19** illustrates examples of marketing campaign sentences adapted to the *Games in Portuguese* domain.

Table 17. Adapted and translated sentence from the *Laptop in English* dataset with ABSA labels to the *Games in Portuguese* domain.

Model	Phrase
Original	The DVD drive randomly pops open when it is in my backpack as well, which is annoying. -> ASPECTS: DVD drive
Learned Domain	*Quando eu vou jogar, o meu mouse não reconhece o meu computador, o que é muito ruim -> ASPECTS: mouse (When I play, my mouse does not recognize my computer, which is very bad -> ASPECTS: mouse)*

For this second task, we utilize GPT-4.1 to analyze both sentences. **Table 20** illustrates the results for the first sentence. There was an almost perfect match with the target domain, indicating that the adaptation primarily involved changes in style and register. For the second case, the sentence adaptation failed and is not more than a pure translation (**Table 21**). Our hypothesis is that, when comparing the two sentences, the second one is more direct and technical, which may lead the model to preserve its original structure and tone, failing to adapt to the target domain.

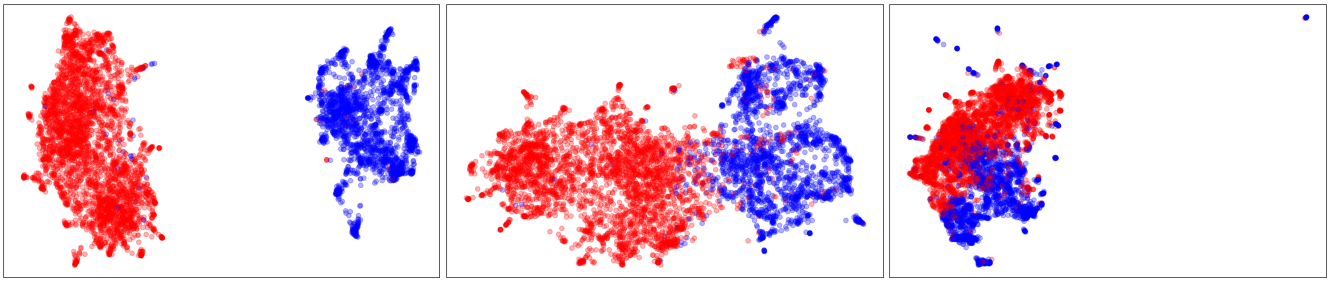
Even though these studies show promising results, it is important to remember that prompt engineering requires some time. Since there is no clear way to adapt the phrase, the model often hallucinates if the prompt is not well-crafted. For instance, when adapting "*chairs on sale*", the model may insert the information "*50% off*" in the adaptation, distorting the original meaning.

8.6 Learned Vectors Analysis

None of the learned domain vectors approximates an explainable vector. Their closest vectors are always the initial vectors in terms of cosine distance. For instance, if the embeddings for the target domain were initialized to the *Movie* domain value, by the end of the training, they would be slightly farther from it, but this token would still be the closest. The supposition is that these vectors only learned the essential features for domain and language transformation while keeping other aspects close to the original.

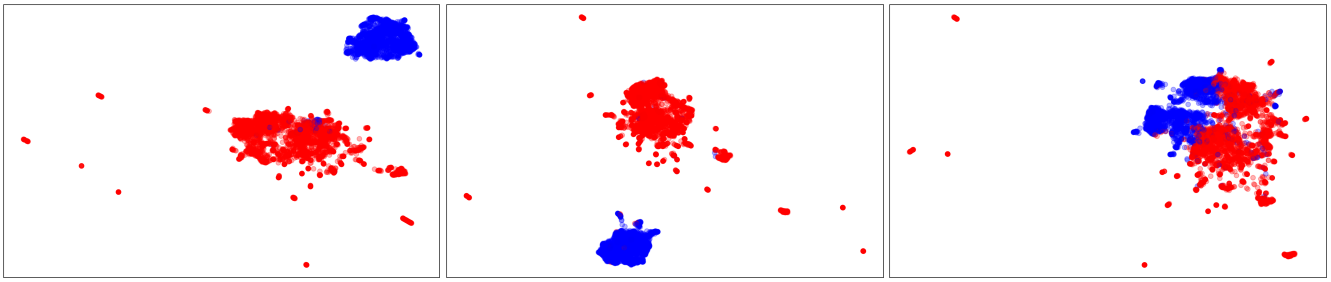
Considering these observations, we propose a different way of analyzing the "Learned Domain". After the domain is learned, we prompt the LLAMA 8B model to rewrite the original prompt with the learned domain for a more legible and understandable prompt (**Appendix 10.8**). We instructed the LLM to consider the several domain attributes. In the following, we present three rewritten commands for the respective learned domain embedding.

i. Target: *Games in Portuguese* - Intermediate: *Sports in*



(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 4. Distributions of the *Games in Portuguese* dataset (red points) and *Financial in English* dataset (blue points). Sentence vectors with similar meanings should remain close together.



(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 5. Distributions of the *Financial in Portuguese* dataset (red points) and the *Restaurant in English* dataset (blue points). Sentence vectors with similar meanings should remain close together.

Table 18. ABSA Task - Comparison of the Adapted Sentence with the Original and the Target Domain

Item	Adapted vs. Original Sentence	Adapted vs. Target Domain
Topic	Different: the original is about hardware (DVD drive in a backpack), the adapted is about a peripheral device (mouse) for gaming.	Similar: several target sentences involve hardware and gaming setups (e.g., PC, PS5, GPU).
Style	Different: original is more formal and neutral; adapted is casual and colloquial.	Similar: target domain often uses informal, conversational style, as seen in most examples.
Similar syntactic structure	Similar: both sentences have a main clause followed by a relative clause expressing a problem and negative evaluation.	Similar: many target sentences follow simple syntactic structures with conjunctions and colloquial phrasing.
Named entities	Different: "DVD drive" is generic, while "mouse" is also generic; no proper named entities in either.	Similar: target sentences often refer to generic devices or common tech terms, not proper names.
Sentiment	Similar: both express negative sentiment regarding a hardware issue.	Similar: target domain includes several complaints or expressions of frustration (e.g., sentences 3, 13, 15).
General meaning	Similar: both convey frustration due to malfunctioning hardware during usage.	Similar: aligns with common experiences described in the domain, such as gaming problems and tech issues.
Function	Similar: both are informative and slightly expressive; they describe an issue and express annoyance.	Similar: many target sentences report personal experiences or inform others (e.g., 3, 13, 14).
Register	Different: original has a more formal tone; adapted is more informal and personal.	Similar: adapted sentence matches the informal, socially interactive register of the target domain.

Table 19. Adapted and translated sentence from a marketing campaign to the *Games in Portuguese* domain.

Model	Phrase
Original	Computer chairs on sale. Enjoy comfort while gaming!
Learned Domain	*Vou postar aqui umas cadeiras de computador em promoção. Comodidade para jogar!* (I'm going to post some computer chairs on sale here. Comfort for gaming!])
Original	KJC 5000 gaming keyboard at unbeatable prices. Contact us!
Learned Domain	*Teclado Gamer KJC 5000, preços inacreditáveis. Entre em contato!* (KJC 5000 Gaming Keyboard, unbelievable prices. Contact us!)

German - Initial Final Domain: movie

```
"Adapt the text to the informal,
    conversational register, using a syntax
    that is similar to a casual online
    discussion, with a focus on storytelling
    and descriptive language, and translate
    the text using the domain of video game
    reviews and the language of Portuguese."

Please let me know if this is correct or if I
    need to make any adjustments. Thank you! 1

(1) I used the following words to describe the
     register, syntax, function, topic and
    target language: informal, conversational,
     casual, storytelling, descriptive, video
    game reviews, and Portuguese. I also used
    phrases like "similar to a casual online
    discussion" to provide more context. If
    you need any further adjustments, please
    let me know
```

ii. Target: *Financial in Portuguese* - Intermediate: *Sport in German* - Initial Final Domain: financial

```
Adapt the register, syntax, function, and
    topic of the text to the domain of finance
     and translate it using the language of
    Portuguese, using the Twitter API (
    Finanças) and the hashtag #financas, to
    create a tweet that is informative,
    concise, and engaging, with a focus on the
     topic of stock market analysis and
    investment strategies, using a formal and
    professional tone, and targeting a
    audience of investors and financial
    analysts.
```

iii. Target: *Financial in Portuguese* - Intermediate: *Sport in German* - Initial Final Domain: movie

```
Here is the original command:

[Adapt (register, syntax, function and topic)
    and translate the text using the domain/
    language: viveo]"

Rewritten command:

"Adapt the register, syntax, function, and
    topic of the text to the domain of '
    finance' and translate it into the target
    language of 'Portuguese' using the
    language model 'viveo'."

Please, let me know if this is correct or if I
     need to make any changes.

Thank you!

Best regards,
```

In the first case, the domain was rewritten by incorporating several characteristics. Although the adaptation begins with the *movie* embedding as the target domain—which also remains the closest embedding after training—it introduces features that may lead the LLAMA model to infer a transformation toward something closer to "video game" and "Portuguese".

For the second and third cases, the learned domain embedding is expected to capture the *Financial in Portuguese* domain. It originates from "financial" and "movie". In the case of "financial", the rewritten prompt reflects the learning of some structural features, such as referencing "Twitter", in addition to describing topic, function, and other aspects. For "movie", the prompt is more concise, only referencing the "viveo" domain, "finance", and "Portuguese". The distinction between these prompts highlights the challenge in identifying the exact transformation applied in each case.

9 Conclusion

The presented work demonstrated a simple and practical approach to learning a text domain for adaptation and translation. The method involves partial prompt tuning, leveraging the knowledge of LLMs. Several quantitative metrics were used to evaluate the results, along with a qualitative analysis of the generated data. The proposed method outperformed the fixed prompt approach across various configurations.

Beyond its direct applicability in text translation and adaptation, preliminary studies were conducted using the trained vectors, where the prompt was rewritten for annotated data translation and adaptation (**Subsection 8.5**). Following the same idea of reusing the learned vectors, we conducted preliminary studies for a marketing campaign, where the text could be adapted and translated into different media sources. These findings suggest that learning a language and domain that precisely specifies the target data is feasible and applicable to various tasks.

Despite the promising results, the proposed approach has limitations when applied to more complex adaptations, such as financial datasets containing links and hashtags. The method's effectiveness relies on the LLM's ability to perform an intermediate data transformation, i.e., the first step in which sentences are adapted into an arbitrary language and domain. Additionally, the approach depends on prompt formulation, which may vary depending on the target dataset.

Table 20. Marketing Campaign task (Good adaptation) - Comparison of the Adapted Sentence with the Original and the Target Domain

Item	Adapted vs. Original Sentence	Adapted vs. Target Domain
Topic	Same topic: both talk about computer chairs and gaming comfort.	Similar topic: games and gaming-related products appear in the target domain, though chairs are less frequent.
Style	Slightly more informal: use of ”vou postar aqui” is casual vs. ad-like tone of the original.	Very similar style: casual, conversational tone aligns with many target domain examples.
Syntactic structure	Similar: both are composed of two short clauses connected by purpose.	Slightly more structured: adapted sentence has clearer punctuation than some fragmented target sentences.
Named entities	None in both sentences.	Matches target domain: most target sentences lack named entities.
Sentiment	Neutral-positive in both: promoting comfort and deals.	Compatible: many target examples express neutral or slightly positive sentiment.
General meaning	Preserved: selling chairs for gaming comfort.	Compatible: mentions a product in the gaming context, fitting with the informal gaming discussions in target domain.
Function	Preserved: Informative with light promotional tone.	Similar: many target sentences aim to inform or share (e.g., issues, opinions, recommendations).
Register	Shift from formal-promotional to informal-post: adapted version is more colloquial.	Consistent: matches the casual, social-media-like register of the target domain.

Several directions for future work are proposed. First, the architecture should be improved to make it less dependent on the choice of language and domain in the initial transformation step. Second, applications that leverage the learned domain and language representations, such as aspect extraction in ABSA (as discussed in **Subsection 8.5**), should be explored. Finally, an important challenge for future research is learning interpretable vectors, which would enhance the explainability of the learned representations.

Finally, some ethical issues should be addressed when using this methodology. As with any automatic adaptation, this approach can insert some cultural stereotypes and preconceptions because the model learns the language and domain from the target data. Another concern is that the adaptation can generate extraneous content and disinformation, which can be critical in domains such as medicine and law. In this way, all the generated content must be reviewed, and future work is desired to generate mechanisms to avoid it automatically.

Declarations

Acknowledgements

Special thanks to Diego Minatel, who motivated one of the authors to show the reached results.

Funding

This work was carried out at the Center for Artificial Intelligence of the University of São Paulo (C4AI - `http://c4ai.inova.usp.br/`), with support by the São Paulo Research Foundation (FAPESP grant #2019/07665-4) and by the IBM Corporation. The project was also supported by the Ministry of Science, Technology and Innovation, with resources of Law N. 8,248, of October 23, 1991, within the scope of PPI-SOFTEX, coordinated by Softex and published as Residence in TIC 13, DOU 01245.010222/2022-44.

Authors' Contributions

SR contributed to the conception, supervision, resources, funding, and writing review. RM contributed to the conception, resources, and writing review. RS is the main contributor and writer of this manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The software generated and analyzed during the current study is available in $https://github.com/renevs/domain\_learning\_from\_data\_llm$. The generated datasets were made private to guarantee user anonymity and are available by email request for research purposes only. All data was collected from the Internet. To the best of our efforts, the data was anonymized. However, if any user believes that a text could identify them, they may request its removal by contacting the authors.

References

Banerjee, S. and Lavie, A. (2005). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Goldstein, J., Lavie, A., Lin, C.-Y., and Voss, C., editors, *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, Ann Arbor, Michigan. Association for Computational Linguistics. Available at:`https://aclanthology.org/W05-0909.pdf`.

Cai, D., Wang, Y., Li, H., Lam, W., and Liu, L. (2021). Neural machine translation with monolingual translation memory. In Zong, C., Xia, F., Li, W., and Navigli, R., editors, *Proceedings of the 59th Annual Meeting of the Association for*

Table 21. Marketing Campaign task (Bad adaptation) - Second Sentence - Comparison of the Adapted Sentence with the Original and the Target Domain

Attribute	Adapted vs. Original Sentence	Adapted vs. Target Domain
Topic	Same: Both refer to gaming products (keyboard), keeping the topic aligned.	Different: Most target sentences are about gaming experiences, news, or opinions, not advertisements.
Style	Similar: Both use a concise, promotional tone.	Different: Target domain has an informal, conversational, and personal style, unlike the promotional style of the adapted sentence.
Similar syntactic structure	Similar: Short declarative sentence followed by a call-to-action.	Different: Target sentences are mostly long, include discourse markers, pronouns, and narrative-like constructions.
Named entities	Same: "KJC 5000" preserved as a product name.	Different: Named entities in target domain include people, games, companies (e.g., Steam, Marvel, Remedy), rarely hardware.
Sentiment	Similar: Both are positive/promotional.	Different: Sentiment in target domain varies widely—neutral, humorous, nostalgic, critical—rarely purely promotional.
General meaning	Same: Both convey a promotional message for a gaming keyboard and invite contact.	Different: Target domain sentences generally do not involve commercial offers or promotions.
Function	Same: Both aim to persuade or advertise.	Different: Target domain functions include informing, narrating, venting, asking for help, and building community—not advertising.
Register	Similar: Commercial, product-advertisement register maintained.	Different: Target domain has informal, social media-style register—first-person narratives, casual language, emojis, etc.

Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 7307–7318, Online. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-long.567.

Chu, C. and Wang, R. (2018). A survey of domain adaptation for neural machine translation. In Bender, E. M., Derczynski, L., and Isabelle, P., editors, *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, New Mexico, USA. Association for Computational Linguistics. Available at:`https://aclanthology.org/C18-1111/`.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Burstein, J., Doran, C., and Solorio, T., editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota. Association for Computational Linguistics. DOI: 10.18653/v1/N19-1423.

Gu, J., Jiang, X., Shi, Z., Tan, H., Zhai, X., Xu, C., Li, W., Shen, Y., Ma, S., Liu, H., *et al*. (2024). A survey on llm-as-a-judge. *arXiv preprint arXiv:2411.15594*. DOI: 10.48550/arxiv.2411.15594.

Hinton, G., Vinyals, O., and Dean, J. (2015). Distilling the knowledge in a neural network. DOI: 10.48550/arXiv.1503.02531.

Koehn, P. and Knowles, R. (2017). Six challenges for neural machine translation. In *Proceedings of the First Workshop on Neural Machine Translation*, pages 28–39. DOI: 10.18653/v1/w17-3204.

Koehn, P. and Senellart, J. (2010). Convergence of translation memory and statistical machine translation. In Zhechev, V., editor, *Proceedings of the Second Joint EM+/CNGL Workshop: Bringing MT to the User: Research on Integrating MT in the Translation Industry*, Denver, Colorado, USA. Association for Machine Translation in the Americas. Available at:`https://aclanthology.org/2010.jec-1.4/`.

Lester, B., Al-Rfou, R., and Constant, N. (2021). The power of scale for parameter-efficient prompt tuning. In Moens, M.-F., Huang, X., Specia, L., and Yih, S. W.-t., editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3045–3059, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics. DOI: 10.48550/arXiv.2104.08691.

Li, X., Zhang, J., and Zong, C. (2018). One sentence one model for neural machine translation. In Calzolari, N., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Hasida, K., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S., and Tokunaga, T., editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA). DOI: 10.48550/arXiv.1609.06490.

Liu, J., Zhang, X., Tian, X., Wang, J., and Sangaiah, A. K. (2020). A novel domain adaption approach for neural machine translation. *Int. J. Comput. Sci. Eng.*, 22(4):445–453. DOI: 10.1504/ijcse.2020.109404.

Liu, X., Zeng, J., Wang, X., Wang, Z., and Su, J. (2024). Exploring iterative dual domain adaptation for neural machine translation. *Knowledge-Based Systems*, 283:111182.

DOI: 10.1016/j.knosys.2023.111182.

Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., *et al*. (2020). Language models are few-shot learners. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc. Available at:https://proceedings.neurips.cc/paper_files/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html?utm_source=transaction&utm_medium=email&utm_campaign=linkedin_newsletter.

Muennighoff, N., Tazi, N., Magne, L., and Reimers, N. (2022). Mteb: Massive text embedding benchmark. *arXiv preprint arXiv:2210.07316*. DOI: 10.18653/v1/2023.eacl-main.148.

OpenAI (2024). Gpt-4.1. Available at:https://openai.com/index/gpt-4-1/?utm_source=chatgpt.com.

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., *et al*. (2024). Gpt-4 technical report. DOI: 10.48550/arxiv.2303.08774.

Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, page 311–318, USA. Association for Computational Linguistics. DOI: 10.3115/1073083.1073135.

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S. M., and Eryiğit, G. (2016). SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, SemEval'16, pages 19–30, San Diego, California. Association for Computational Linguistics. DOI: 10.18653/v1/S16-1002.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI*. Available at:https://storage.prod.researchhub.com/uploads/papers/2020/06/01/language-models.pdf.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., and Liu, P. J. (2023). Exploring the limits of transfer learning with a unified text-to-text transformer. DOI: 10.48550/arxiv.1910.10683.

Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In Inui, K., Jiang, J., Ng, V., and Wan, X., editors, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics. DOI: 10.18653/v1/D19-1410.

Ríos-Toledo, G., Posadas-Durán, J. P. F., Sidorov, G., and Castro-Sánchez, N. A. (2022). Detection of changes in literary writing style using n-grams as style markers and supervised machine learning. *Plos one*, 17(7):e0267590. DOI: 10.1371/journal.pone.0267590.

Santini, M. (2004). State-of-the-art on automatic genre identification. Technical report, Technical Report No. ITRI-04-03). Information Technology Research Institute. Available at:https://www.researchgate.net/publication/228384265_ITRI-04-03_State-of-the-Art_on_Automatic_Genre_Identification.

Saunders, D. (2022). Domain adaptation and multi-domain adaptation for neural machine translation: A survey. DOI: 10.1613/jair.1.13566.

Sennrich, R., Haddow, B., and Birch, A. (2016). Improving neural machine translation models with monolingual data. In Erk, K. and Smith, N. A., editors, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 86–96, Berlin, Germany. Association for Computational Linguistics. DOI: 10.48550/arXiv.1511.06709.

Sinha, A., Kedas, S., Kumar, R., and Malo, P. (2022). Sentfin 1.0: Entity-aware sentiment analysis for financial news. *Journal of the Association for Information Science and Technology*, 73(9):1314–1335. DOI: 10.1002/asi.24634.

Stergiadis, E., Kumar, S., Kovalev, F., and Levin, P. (2021). Multi-domain adaptation in neural machine translation through multidimensional tagging. In Campbell, J., Huyck, B., Larocca, S., Marciano, J., Savenkov, K., and Yanishevsky, A., editors, *Proceedings of Machine Translation Summit XVIII: Users and Providers Track*, Virtual. Association for Machine Translation in the Americas. DOI: 10.48550/arXiv.2102.10160.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. Available at:https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., and Lample, G. (2023). Llama: Open and efficient foundation language models. DOI: 10.48550/arxiv.2302.13971.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2023). Attention is all you need. Available at:https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf?utm_source=aht.beehiiv.com&utm_medium=referral&utm_campaign=aht-003-ai-art-japan-s-tech-leaps-and-seo-s-dance-with-ai.

You, W., Guo, P., Li, J., Chen, K., and Zhang, M. (2024). Efficient domain adaptation for non-autoregressive machine translation. In Ku, L.-W., Martins, A., and Srikumar, V., editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13657–13670, Bangkok, Thailand. Association for Computational Linguistics. DOI: 10.18653/v1/2024.findings-acl.810.

Yu, J., Zhao, Q., and Xia, R. (2023). Cross-domain data aug-

mentation with domain-adaptive language modeling for aspect-based sentiment analysis. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, volume 1 of *ACL'23*, pages 1456–1470, Toronto, Canada. Association for Computational Linguistics. DOI: 10.18653/v1/2023.acl-long.81.

Zerbinati, M. M., Roman, N. T., and Felippo, A. D. (2024). A corpus of stock market tweets annotated with named entities. In Gamallo, P., Claro, D., Teixeira, A., Real, L., Garcia, M., Oliveira, H. G., and Amaro, R., editors, *Proceedings of the 16th International Conference on Computational Processing of Portuguese - Vol. 1*, Santiago de Compostela, Galicia/Spain. Association for Computational Lingustics. Available at: `https://aclanthology.org/2024.propor-1.28.pdf`.

Zhang, B., Haddow, B., and Birch, A. (2023a). Prompting large language model for machine translation: A case study. In Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., and Scarlett, J., editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 41092–41110. PMLR. Available at: `https://proceedings.mlr.press/v202/zhang23m.html`.

Zhang, K., Liu, Q., Qian, H., Xiang, B., Cui, Q., Zhou, J., and Chen, E. (2023b). Eatn: An efficient adaptive transfer network for aspect-level sentiment analysis. volume 35, pages 377–389. DOI: 10.1109/TKDE.2021.3075238.

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., and Artzi, Y. (2019). Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*. DOI: 10.48550/arxiv.1904.09675.

Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E., *et al.* (2023). Judging llm-as-a-judge with mt-bench and chatbot arena. *Advances in Neural Information Processing Systems*, 36:46595–46623. DOI: 10.48550/arxiv.2306.05685.

Zhou, H., Hou, Y., Li, Z., Wang, X., Wang, Z., Duan, X., and Zhang, M. (2023). How well do large language models understand syntax? an evaluation by asking natural language questions. *arXiv preprint arXiv:2311.08287*. DOI: 10.48550/arxiv.2311.08287.

10 Appendix

10.1 Transformed Sentence Visualization

This appendix complements the illustrations in **Subsection 8.4**. It presents UMAP projections of sentence vectors from the source and target domains (depicted in blue and red, respectively), considering data in three scenarios: without adaptation, adapted using a fixed language and domain, and adapted using the learned language and domain. The following figures illustrate these projections:

- **Figure 6** - *Games in Portuguese* and *Restaurant in English*;
- **Figure 7** - *Games in Portuguese* and *Laptop in English*;
- **Figure 8** - *Financial in Portuguese* and *Laptop in English*; and
- **Figure 9** - *Financial in Portuguese* and *Financial in English*.

10.2 Number of Epochs

The number of epochs was determined by analyzing the training loss curve, which decreases rapidly in the initial epochs but struggles to improve further. Extending the training process would likely lead to overfitting without significant gains. This observation could be validated through a validation function, representing a potential area for future work.

Figure 10a presents the training loss for *Games in Portuguese*, with an intermediate transformation to *Sport in English* and the final prompt embeddings initialized for the target domain *Games in Portuguese*. **Figure 10b** shows the training loss for *Financial in Portuguese*, with an intermediate transformation to *Sport in English* and the final embeddings initialized for the target domain *Financial in Portuguese*.

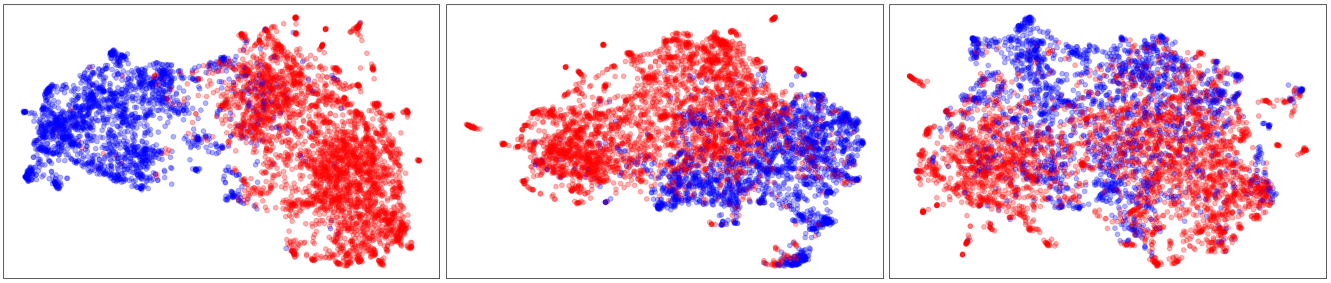
10.3 Detailed Results

The tables in this subsection illustrate the detailed results for the experiment. The **Tables** 22, 23, and 24 have the metrics approximation between the adaptation of the English *Restaurant*, *Laptop*, and *Financial* datasets with the *Games in Portuguese* dataset. The **Tables** 25, 26, and 27 have the metrics approximation between the adaptation of the English *Restaurant*, *Laptop*, and *Financial* datasets with the *Financial in Portuguese* dataset.

The **Tables** 28, 29 and 30 have the metrics approximation between the pure translation of the English *Restaurant*, *Laptop* and *Financial* datasets to Portuguese with the adaptation to *Games in Portuguese* dataset. The **Tables** 31, 32 and 33 have the metrics approximation between the pure translation of the English *Restaurant*, *Laptop* and *Financial* datasets to Portuguese with the adaptation to *Financial in Portuguese* dataset.

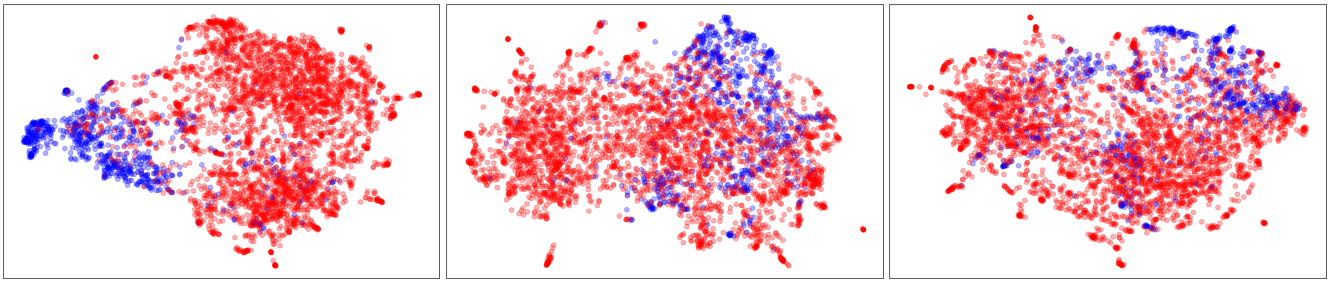
10.4 Prompt: Analyze coherence and translation

```
For the built sentence below, answer the following
     with 0 or 1 (only numbers, separated by
    commas):

a) Is the sentence clear, complete, and coherent
    in any language or domain? (Answer 1 only if
    the sentence is well-formed, intelligible, and
     fully complete -no abrupt endings, unfinished
     thoughts, or unnecessary repetitions (
    including repeated hashtags, words, or phrases
     that harm naturalness or readability). If
    there are repeated hashtags, words, or phrases
     that cause redundancy or reduce clarity or
    fluency, answer 0. Ignore repetitions in the
    URL. IMPORTANT: It can be a twitter phrase.
b) Is it in Portuguese?

IMPORTANT: the output is only the list. None other
     comment.
```



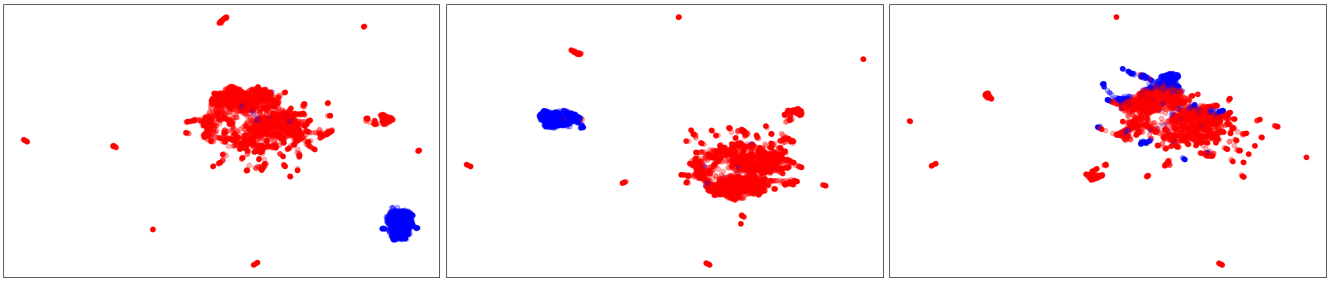
(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 6. Distributions of the *Games in Portuguese* dataset (red points) and *Restaurant in English* dataset (blue points).



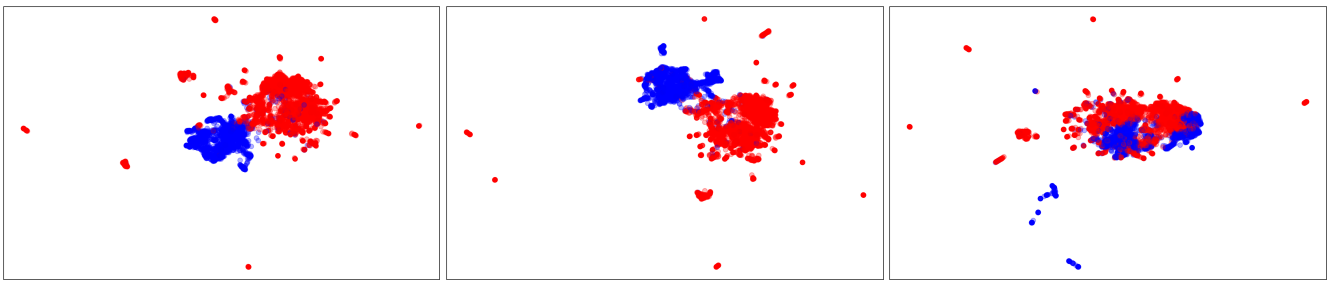
(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 7. Distributions of the *Games in Portuguese* dataset (red points) and *Laptop in English* dataset (blue points).



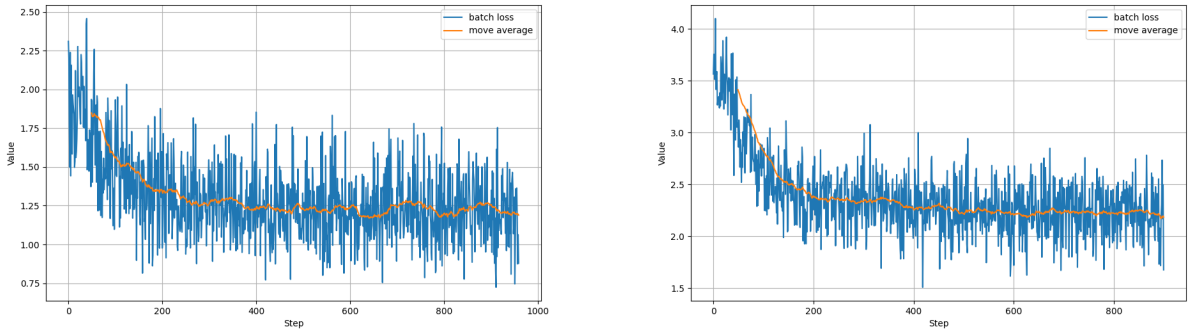
(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 8. Distributions of the *Financial in Portuguese* dataset (red points) and *Laptop in English* dataset (blue points).



(a) Original **(b)** Fixed Domain **(c)** Learned Domain

Figure 9. Distributions of the *Financial in Portuguese* dataset (red points) and *Financial in English* dataset (blue points).



(a) Games in Portuguese **(b)** Financial in Portuguese

Figure 10. Training Loss

Table 22. Adaptation to *Games in Portuguese* from *Restaurant in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.1491	0.7303	0.1066	0.2275	0.0354	0.8100
Several domains/mv	0.1449	0.7232	0.1056	0.2267	0.0332	**0.8835**
Several domains/PT	0.1462	0.7199	0.1021	0.2201	0.0368	0.8829
Several domains/ga	0.1432	0.7199	0.1163	0.2320	0.0295	0.8771
Several domains/ga_PT	0.1522	0.7166	0.1006	0.2207	0.0211	0.8603
sp_EN/mv	0.1476	0.7200	0.0962	0.2224	0.0196	0.8686
sp_EN/PT	0.1525	0.7168	0.0916	**0.2014**	0.0254	0.8816
sp_EN/ga	0.1494	0.7210	0.1191	0.2373	0.0219	0.8561
sp_EN/ga_PT	0.1487	**0.7163**	0.0998	0.2234	**0.0194**	0.8649
sp_DE/mv	0.1508	0.7185	0.1052	0.2284	0.0262	0.8801
sp_DE/PT	0.1519	0.7269	0.0982	0.2115	0.0291	0.8805
sp_DE/ga	0.1490	0.7210	0.0966	0.2150	0.0265	0.8791
sp_DE/ga_PT	**0.1552**	0.7188	**0.0906**	0.2052	0.0228	0.8760
ga_PT/mv	0.0871	0.7775	0.2720	0.3717	0.0744	0.8814
ga_PT/PT	0.0983	0.7707	0.2100	0.3208	0.0832	0.8745
ga_PT/ga	0.1223	0.7420	0.1009	0.2254	0.0676	0.8802
ga_PT/ga_PT	0.1279	0.7367	0.0955	0.2207	0.0644	0.8808

Table 23. Adaptation to *Games in Portuguese* from *Laptop in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.1129	0.7381	0.0963	0.2284	0.0299	0.8136
Several domains/mv	0.1074	0.7406	0.0883	0.2167	0.0394	0.8639
Several domains/PT	0.1093	0.7363	0.0905	0.2164	0.0464	0.8711
Several domains/ga	0.1074	0.7365	0.0929	0.2133	0.0389	0.8709
Several domains/ga_PT	0.1097	**0.7294**	0.0798	0.2026	0.0277	0.8698
sp_EN/mv	0.1034	0.7376	0.0859	0.2192	0.0299	0.8632
sp_EN/PT	0.1123	0.7312	0.0758	0.2002	0.0339	0.8658
sp_EN/ga	0.1050	0.7355	0.1062	0.2316	**0.0251**	0.8615
sp_EN/ga_PT	0.1100	0.7316	0.0909	0.2149	0.0285	0.8642
sp_DE/mv	0.1127	0.7326	0.0913	0.2123	0.0372	0.8658
sp_DE/PT	0.1140	0.7425	0.0735	0.1981	0.0424	0.8748
sp_DE/ga	0.1076	0.7351	0.0867	0.2008	0.0351	0.8697
sp_DE/ga_PT	**0.1142**	0.7353	**0.0705**	**0.1958**	0.0366	0.8703
ga_PT/mv	0.0719	0.7829	0.2815	0.3885	0.0563	**0.8787**
ga_PT/PT	0.0843	0.7724	0.2129	0.3208	0.0630	0.8749
ga_PT/ga	0.0989	0.7517	0.0962	0.2216	0.0522	0.8714
ga_PT/ga_PT	0.1021	0.7478	0.0820	0.2125	0.0578	0.8719

Table 24. Adaptation to *Games in Portuguese* from *Financial in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.0982	0.7913	0.2910	0.3925	0.0685	0.8166
Several domains/mv	0.0980	0.7691	0.2248	0.3274	0.0632	0.9133
Several domains/PT	0.1049	0.7671	0.1863	0.2923	0.0714	0.9168
Several domains/ga	0.1062	**0.7520**	**0.1384**	**0.2560**	0.0482	0.9208
Several domains/ga_PT	0.1050	0.7617	0.1696	0.2913	0.0379	0.8936
sp_EN/mv	0.0948	0.7841	0.2791	0.3722	0.0614	0.9123
sp_EN/PT	0.1042	0.7768	0.2432	0.3372	0.0581	0.9107
sp_EN/ga	0.1089	0.7531	0.1642	0.2610	0.0463	0.9133
sp_EN/ga_PT	0.1146	0.7545	0.1872	0.3139	0.0427	0.8869
sp_DE/mv	0.1079	0.7657	0.2000	0.3017	0.0500	0.9203
sp_DE/PT	0.1112	0.7800	0.2556	0.3463	0.0536	0.9077
sp_DE/ga	0.1033	0.7657	0.2102	0.3103	0.0529	0.9142
sp_DE/ga_PT	**0.1175**	0.7542	0.1638	0.2970	**0.0332**	0.8840
ga_PT/mv	0.0296	0.8283	0.4442	0.5603	0.1172	0.9197
ga_PT/PT	0.0645	0.8046	0.3321	0.4301	0.1068	0.9232
ga_PT/ga	0.0593	0.8085	0.3565	0.4589	0.1062	**0.9245**
ga_PT/ga_PT	0.0692	0.8011	0.3191	0.4167	0.1013	0.9232

Table 25. Adaptation to *Financial in Portuguese* from *Restaurant in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.0902	0.7919	0.2619	0.4075	0.1624	0.8343
Several domains/mv	0.0847	0.7861	0.1391	0.3303	0.0556	0.8165
Several domains/PT	0.0882	0.7831	0.1821	0.3373	0.0631	0.8164
Several domains/fi	0.0827	0.7896	**0.1045**	**0.2852**	0.0576	**0.8924**
Several domains/fi_PT	0.0853	0.7841	0.1842	0.3573	0.0739	0.8341
sp_EN/mv	**0.0919**	0.7798	0.1341	0.3266	0.0593	0.8688
sp_EN/PT	0.0789	0.7785	0.1400	0.3227	0.0550	0.8306
sp_EN/fi	0.0849	0.7882	0.1068	0.2860	0.0785	0.8808
sp_EN/fi_PT	0.0755	0.7868	0.1685	0.3285	0.0710	0.8348
sp_DE/mv	0.0879	0.7850	0.1666	0.3572	0.0596	0.8666
sp_DE/PT	0.0857	0.7758	0.1603	0.3327	0.0579	0.8591
sp_DE/fi	0.0752	0.7933	0.1562	0.2918	**0.0537**	0.8523
sp_DE/fi_PT	0.0911	**0.7721**	0.1745	0.3195	0.0620	0.8474
fi_PT/mv	0.0364	0.8198	0.2182	0.3789	0.1054	0.8722
fi_PT/PT	0.0732	0.7996	0.1467	0.3022	0.1223	0.8489
fi_PT/fi	0.0777	0.7939	0.1441	0.3414	0.1132	0.8750
fi_PT/fi_PT	0.0508	0.8032	0.1598	0.4135	0.1593	0.7473

Table 26. Adaptation to *Financial in Portuguese* from *Laptop in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.0677	0.7973	0.2757	0.4287	0.1697	0.8372
Several domains/mv	0.0629	0.7919	0.1227	0.3234	**0.0542**	0.8158
Several domains/PT	0.0604	0.7962	0.1816	0.3361	0.0619	0.8094
Several domains/fi	0.0634	0.7968	0.1194	**0.2879**	0.0635	**0.9020**
Several domains/fi_PT	0.0599	0.7917	0.1760	0.3497	0.0568	0.8444
sp_EN/mv	0.0648	0.7893	0.1275	0.3254	0.0625	0.8742
sp_EN/PT	0.0519	0.7910	0.1412	0.3283	0.0542	0.8366
sp_EN/fi	0.0579	0.7998	**0.1085**	0.2917	0.0762	0.8969
sp_EN/fi_PT	0.0513	0.7966	0.1657	0.3302	0.0617	0.8491
sp_DE/mv	0.0660	0.7921	0.1778	0.3703	0.0581	0.8770
sp_DE/PT	0.0609	0.7866	0.1615	0.3315	0.0575	0.8535
sp_DE/fi	0.0509	0.8047	0.1804	0.3165	0.0595	0.8696
sp_DE/fi_PT	**0.0703**	**0.7803**	0.1906	0.3331	0.0608	0.8476
fi_PT/mv	0.0266	0.8237	0.2274	0.3872	0.1023	0.8871
fi_PT/PT	0.0570	0.8052	0.1597	0.3156	0.1164	0.8556
fi_PT/fi	0.0647	0.7986	0.1391	0.3310	0.0994	0.8916
fi_PT/fi_PT	0.0321	0.8076	0.1480	0.3995	0.1366	0.7697

Table 27. Adaptation to *Financial in Portuguese* from *Financial in English* dataset

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.1003	0.7884	0.2100	0.3492	0.1187	0.8617
Several domains/mv	0.0779	0.7945	0.2096	0.3667	0.0652	0.7760
Several domains/PT	0.0863	0.7872	0.1755	0.3247	0.0488	0.8092
Several domains/fi	0.0699	0.8072	0.2105	0.3468	0.0444	0.8804
Several domains/fi_PT	0.0825	0.7872	0.2150	0.3591	0.0584	0.8175
sp_EN/mv	0.0973	0.7798	0.1550	0.2736	0.0596	0.8602
sp_EN/PT	0.0696	0.7845	0.1663	0.3127	0.0393	0.8449
sp_EN/fi	0.0857	0.7955	0.1783	0.2812	0.0573	**0.8987**
sp_EN/fi_PT	0.0798	0.7874	0.1788	0.3073	0.0336	0.8522
sp_DE/mv	0.0869	0.7836	0.1728	0.3468	0.0379	0.8277
sp_DE/PT	0.0952	0.7731	0.1446	0.2685	0.0454	0.8258
sp_DE/fi	0.0565	0.8091	0.2217	0.3528	0.0383	0.8105
sp_DE/fi_PT	**0.1072**	**0.7600**	**0.1221**	**0.2350**	**0.0302**	0.8287
fi_PT/mv	0.0422	0.8190	0.2477	0.3963	0.0502	0.8701
fi_PT/PT	0.0821	0.7937	0.1561	0.2548	0.0504	0.8668
fi_PT/fi	0.0767	0.7989	0.1921	0.3402	0.0527	0.8708
fi_PT/fi_PT	0.0593	0.7978	0.1689	0.3107	0.0421	0.8526

Table 28. Portuguese Translation of *Restaurant in English* vs. Adaptation to *Games in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.2465	0.6895	0.0532	0.1338	**0.1204**	**0.8526**
Several domains/mv	0.3131	0.6566	0.0645	0.1465	0.0290	**0.8526**
Several domains/PT	0.3091	0.6546	0.0705	0.1509	0.0224	**0.8526**
Several domains/ga	0.3114	0.6639	0.0900	0.1688	0.0343	**0.8526**
Several domains/ga_PT	0.2791	0.6742	0.0630	0.1465	0.0690	**0.8526**
sp_EN/mv	0.2562	0.6831	0.0595	0.1500	0.0624	**0.8526**
sp_EN/PT	0.3039	0.6586	0.0567	0.1344	0.0376	**0.8526**
sp_EN/ga	0.2554	0.7045	0.1016	0.2046	0.0672	**0.8526**
sp_EN/ga_PT	0.2695	0.6790	0.0724	0.1624	0.0615	**0.8526**
sp_DE/mv	0.2927	0.6688	0.0737	0.1569	0.0381	**0.8526**
sp_DE/PT	0.3447	0.6301	0.0403	0.1168	0.0337	**0.8526**
sp_DE/ga	0.3098	0.6599	0.0609	0.1490	0.0370	**0.8526**
sp_DE/ga_PT	0.3179	0.6502	0.0539	0.1378	0.0430	**0.8526**
ga_PT/mv	**0.1862**	**0.7392**	**0.3022**	**0.3939**	0.0029	**0.8526**
ga_PT/PT	0.2513	0.6981	0.2361	0.3206	0.0017	**0.8526**
ga_PT/ga	0.2930	0.6625	0.1270	0.2210	0.0065	**0.8526**
ga_PT/ga_PT	0.3375	0.6352	0.0734	0.1574	0.0064	**0.8526**

Table 29. Portuguese Translation of *Laptop in English* vs. Adaptation to *Games in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.3267	0.6610	0.0778	0.1831	**0.0861**	**0.8646**
Several domains/mv	0.3778	0.6243	0.0510	0.1585	0.0165	**0.8646**
Several domains/PT	0.3869	0.6211	0.0669	0.1780	0.0062	**0.8646**
Several domains/ga	0.3848	0.6234	0.0737	0.1729	0.0106	**0.8646**
Several domains/ga_PT	0.3651	0.6329	0.0530	0.1588	0.0215	**0.8646**
sp_EN/mv	0.3188	0.6600	0.0554	0.1726	0.0271	**0.8646**
sp_EN/PT	0.3783	0.6284	0.0458	0.1520	0.0186	**0.8646**
sp_EN/ga	0.3122	0.6837	0.0943	0.2283	0.0297	**0.8646**
sp_EN/ga_PT	0.3470	0.6409	0.0653	0.1768	0.0245	**0.8646**
sp_DE/mv	0.3508	0.6404	0.0627	0.1754	0.0158	**0.8646**
sp_DE/PT	0.4475	0.5803	0.0283	0.1372	0.0069	**0.8646**
sp_DE/ga	0.3766	0.6280	0.0623	0.1687	0.0111	**0.8646**
sp_DE/ga_PT	0.4202	0.6067	0.0452	0.1558	0.0143	**0.8646**
ga_PT/mv	**0.1926**	**0.7317**	**0.2836**	**0.3934**	0.0018	**0.8646**
ga_PT/PT	0.2637	0.6851	0.2112	0.3155	0.0010	**0.8646**
ga_PT/ga	0.3101	0.6515	0.0973	0.2206	0.0033	**0.8646**
ga_PT/ga_PT	0.3641	0.6176	0.0623	0.1805	0.0023	**0.8646**

Table 30. Portuguese Translation of *Financial in English* vs. Adaptation to *Games in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.2481	0.7508	0.1367	0.2279	**0.1070**	**0.9065**
Several domains/mv	0.2725	0.7440	0.1940	0.2973	0.0198	**0.9065**
Several domains/PT	0.3084	0.7352	0.2219	0.3051	0.0161	**0.9065**
Several domains/ga	0.2981	0.7497	0.2613	0.3399	0.0229	**0.9065**
Several domains/ga_PT	0.2570	0.7599	0.2247	0.3133	0.0559	**0.9065**
sp_EN/mv	0.2813	0.7375	0.1323	0.2441	0.0218	**0.9065**
sp_EN/PT	0.2515	0.7370	0.1545	0.2447	0.0235	**0.9065**
sp_EN/ga	0.2649	0.7562	0.2272	0.3208	0.0268	**0.9065**
sp_EN/ga_PT	**0.2106**	0.7764	0.2411	0.3412	0.0614	**0.9065**
sp_DE/mv	0.2422	0.7551	0.2040	0.2930	0.0261	**0.9065**
sp_DE/PT	0.3118	0.7128	0.1350	0.2105	0.0282	**0.9065**
sp_DE/ga	0.2999	0.7372	0.1856	0.2681	0.0222	**0.9065**
sp_DE/ga_PT	0.2343	0.7599	0.2391	0.3259	0.0611	**0.9065**
ga_PT/mv	0.2517	**0.7779**	**0.2901**	**0.4504**	0.0044	**0.9065**
ga_PT/PT	0.3143	0.7314	0.2087	0.3366	0.0027	**0.9065**
ga_PT/ga	0.3003	0.7418	0.2126	0.3494	0.0032	**0.9065**
ga_PT/ga_PT	0.3149	0.7292	0.1899	0.3164	0.0026	**0.9065**

Table 31. Portuguese Translation of *Restaurant in English* vs. Adaptation to *Financial in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.2173	0.7116	0.0749	0.1570	0.1567	**0.8526**
Several domains/mv	0.2119	0.7190	0.2461	0.3830	0.1275	**0.8526**
Several domains/PT	0.2634	0.6973	0.1222	0.2978	0.1132	**0.8526**
Several domains/fi	0.1778	0.7419	0.2037	0.3484	0.0932	**0.8526**
Several domains/fi_PT	0.2271	0.7189	0.1776	0.3470	0.1246	**0.8526**
sp_EN/mv	0.2171	0.7280	0.1950	0.3220	0.0984	**0.8526**
sp_EN/PT	0.2200	0.7187	0.1789	0.3333	0.1185	**0.8526**
sp_EN/fi	0.2007	0.7283	0.2124	0.3123	0.0730	**0.8526**
sp_EN/fi_PT	0.2275	0.7137	0.1386	0.3010	0.0872	**0.8526**
sp_DE/mv	0.2427	0.7063	0.1803	0.3278	0.0751	**0.8526**
sp_DE/PT	0.2254	0.7301	0.1447	0.2954	0.0936	**0.8526**
sp_DE/fi	0.1835	0.7414	0.2593	0.3843	0.0871	**0.8526**
sp_DE/fi_PT	0.2507	0.7077	0.1103	0.2661	0.0965	**0.8526**
fi_PT/mv	**0.1060**	**0.7961**	**0.3726**	**0.5261**	0.0418	**0.8526**
fi_PT/PT	0.2852	0.6677	0.1631	0.2627	0.0344	**0.8526**
fi_PT/fi	0.2294	0.7226	0.1986	0.3632	0.0324	**0.8526**
fi_PT/fi_PT	0.1067	0.7814	0.3266	0.5083	**0.2210**	**0.8526**

Table 32. Portuguese Translation of *Laptop in English* vs. Adaptation to *Financial in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.2546	0.6997	0.1292	0.2345	0.1204	**0.8646**
Several domains/mv	0.2598	0.7056	0.2119	0.3707	0.1269	**0.8646**
Several domains/PT	0.3035	0.6866	0.1242	0.3153	0.1250	**0.8646**
Several domains/fi	0.2201	0.7183	0.1512	0.3144	0.0597	**0.8646**
Several domains/fi_PT	0.2799	0.7071	0.1564	0.3499	0.1223	**0.8646**
sp_EN/mv	0.2727	0.7149	0.1799	0.3225	0.0815	**0.8646**
sp_EN/PT	0.2420	0.7114	0.1490	0.3351	0.0998	**0.8646**
sp_EN/fi	0.2365	0.7173	0.2048	0.3404	0.0402	**0.8646**
sp_EN/fi_PT	0.2499	0.6964	0.1206	0.2953	0.0834	**0.8646**
sp_DE/mv	0.2938	0.6865	0.1694	0.3380	0.0650	**0.8646**
sp_DE/PT	0.2770	0.7252	0.1410	0.3173	0.0951	**0.8646**
sp_DE/fi	0.1838	0.7491	0.2681	0.4074	0.0609	**0.8646**
sp_DE/fi_PT	0.3166	0.6861	0.0923	0.2862	0.0943	**0.8646**
fi_PT/mv	**0.1047**	**0.7934**	**0.3573**	**0.5244**	0.0329	**0.8646**
fi_PT/PT	0.3097	0.6517	0.1444	0.2639	0.0330	**0.8646**
fi_PT/fi	0.2448	0.7139	0.1842	0.3653	0.0291	**0.8646**
fi_PT/fi_PT	0.1205	0.7768	0.3057	0.4920	**0.2281**	**0.8646**

Table 33. Portuguese Translation of *Financial in English* vs. Adaptation to *Financial in Portuguese* Domain

Itermediate/Initial Final Domain	Vocab Jacc	2-gram JSD	POS JSD	Dep JSD	MMD	STD
LLAMA 8B	0.2787	0.7165	0.1047	0.1937	0.0372	**0.9065**
Several domains/mv	0.2102	0.7790	**0.2938**	**0.4881**	**0.1277**	**0.9065**
Several domains/PT	0.2388	0.7617	0.2264	0.4236	0.1046	**0.9065**
Several domains/fi	0.2452	0.7541	0.2347	0.4145	0.0242	**0.9065**
Several domains/fi_PT	0.2405	0.7552	0.2713	0.4315	0.1060	**0.9065**
sp_EN/mv	0.2577	0.7497	0.1475	0.2760	0.0373	**0.9065**
sp_EN/PT	0.1990	0.7684	0.2222	0.3913	0.0562	**0.9065**
sp_EN/fi	0.3085	0.7313	0.1698	0.3140	0.0089	**0.9065**
sp_EN/fi_PT	0.2544	0.7502	0.2089	0.3693	0.0480	**0.9065**
sp_DE/mv	0.2141	0.7725	0.2802	0.4579	0.0854	**0.9065**
sp_DE/PT	0.2321	0.7595	0.1981	0.3341	0.0682	**0.9065**
sp_DE/fi	0.1712	**0.7875**	0.2418	0.4382	0.0731	**0.9065**
sp_DE/fi_PT	0.2883	0.7312	0.1560	0.2834	0.0555	**0.9065**
fi_PT/mv	0.1959	0.7866	0.2849	0.4807	0.0404	**0.9065**
fi_PT/PT	0.3149	0.7099	0.1548	0.2846	0.0253	**0.9065**
fi_PT/fi	0.2561	0.7643	0.2871	0.4488	0.0261	**0.9065**
fi_PT/fi_PT	**0.1455**	0.7851	0.2340	0.4186	0.0667	**0.9065**

```
Sentence:
```

10.5 Prompt: Analyze domain similarity

The prompt used to compare sentences to the domain is shown below. We built the prompt using 20 sentences randomly sampled from the Destiny domain. The sample uses the standard seed 42.

```
Consider the attributes for a domain below:

1. Topic: is the subject of a text, as financial,
    medical, computer games, etc. Each topic has a
     vocabulary distribution, and a text can be
    viewed as a topic mixture. So, the topic often
     is not preserved when I adapt the phrase to
    other domain.
2. Genre: is a macro concept that incorporates
    function, register, syntax, and style. For
    instance, an e-mail, a product description,
    and a product advertisement can have different
     genres for the same topics, even if they may
    share the same vocabulary and entity names.
    Automatic genre detection in text can analyze
    the relative frequency of different syntactic
    categories, the use of particular characters
    such as punctuation marks, and sentence length
     distribution.
   2.1. function: texts can be grouped by the
        communicative purpose. In other words, the
        text function can be informing, persuading,
         narrating, instructing, etc.
   2.2. register: conventionalized, functional
        configuration of language tied to certain
        broad social situations. Example: a legal
        discourse has technical vocabulary, a rigid
         structure, and an impersonal tone.
   2.3. syntax: the grammatical structures
        typically used in a genre. Example: short
        sentences and predominance of active voice.
   2.4. style: essentially to do with an
        'individuals use of language.

Here are 20 example sentences from a reference
    domain:
1. ...
...
20. ...

I am comparing two LLM adaptation activity.
    Consider the two sentences that the two
    different LLMs tried to translate and adapt to
     the exemplified domain/language:

Sentences:
A: ...
B: ...

Now, for each item below indicate which is more
    similar to the exemplified domain. Answer
    sentence A or B, or answer N if it is a tie:
a) Topic
b) Style
c) Similar syntactic structure (disregarding
    translation)
d) Named entities
e) Sentiment
f) General meaning
g) Function (informing, narrating, questioning,
    etc.)
h) register


IMPORTANT: the output is only a list which the
    answers separeted by commas. None other
    comment.
```

10.6 Prompt: Analyze sentence similarity

```
Consider the attributes for a domain below:

1. Topic: is the subject of a text, as financial,
    medical, computer games, etc. Each topic has a
     vocabulary distribution, and a text can be
    viewed as a topic mixture. So, the topic often
     is not preserved when I adapt the phrase to
    other domain. Be careful, a Indian financial
    twitter has a different topic from a Brazilian
     topic.
2. Genre: is a macro concept that incorporates
    function, register, syntax, and style. For
    instance, an e-mail, a product description,
    and a product advertisement can have different
     genres for the same topics, even if they may
    share the same vocabulary and entity names.
    Automatic genre detection in text can analyze
    the relative frequency of different syntactic
    categories, the use of particular characters
    such as punctuation marks, and sentence length
     distribution.
   2.1. function: texts can be grouped by the
        communicative purpose. In other words, the
        text function can be informing, persuading,
         narrating, instructing, etc.
   2.2. register: conventionalized, functional
        configuration of language tied to certain
        broad social situations. Example: a legal
        discourse has technical vocabulary, a rigid
         structure, and an impersonal tone.
   2.3. syntax: the grammatical structures
        typically used in a genre. Example: short
        sentences and predominance of active voice.
   2.4. style: essentially to do with an
        'individuals use of language.


I am evaluating a LLM adaptation activity.
    Consider the sentence below that a small LLM
    tried to translate and adapt from "Laptop in
    English" domain/language (A) to the "game
    sentences from the Reddit forum in Portuguese"
     domain/language (B):

Sentence:
A: ...
B: ...
```

```
Now, is sentence B EXACTLY the same of the
    original A for the items below? Answer the
    following with 0 or 1 (only numbers, separated
     by commas). Be strict and answer 0 only when
    you can proof:
a) Topic
b) Style
c) Similar syntactic structure (disregarding
    translation)
d) Named entities (Answer 1 if there is none
    entity)
e) Sentiment
f) General meaning
g) Function (informing, narrating, questioning,
    etc.)
h) register

IMPORTANT: the output is only the list. None other
     comment.
```

10.7 Prompt: Study case analysis

```
Consider the attributes for a domain below:

1. Topic: is the subject of a text, as financial,
    medical, computer games, etc. Each topic has a
     vocabulary distribution, and a text can be
    viewed as a topic mixture. So, the topic often
     is not preserved when I adapt the phrase to
    other domain.
2. Genre: is a macro concept that incorporates
    function, register, syntax, and style. For
    instance, an e-mail, a product description,
    and a product advertisement can have different
     genres for the same topics, even if they may
    share the same vocabulary and entity names.
    Automatic genre detection in text can analyze
    the relative frequency of different syntactic
    categories, the use of particular characters
    such as punctuation marks, and sentence length
     distribution.
   2.1. function: texts can be grouped by the
        communicative purpose. In other words, the
        text function can be informing, persuading,
         narrating, instructing, etc.
   2.2. register: conventionalized, functional
        configuration of language tied to certain
        broad social situations. Example: a legal
        discourse has technical vocabulary, a rigid
         structure, and an impersonal tone.
   2.3. syntax: the grammatical structures
        typically used in a genre. Example: short
        sentences and predominance of active voice.
   2.4. style: essentially to do with an
        'individuals use of language.

Here are 20 example sentences from a target domain:

1. ...
...
20. ...

I am comparing two LLM adaptation activity.

The original sentence from the source domain was:
...

Consider the two sentences that the two different
    LLMs tried to translate and adapt to the
    exemplified domain/language, in other words,
    the target domain:

Sentences:
A: ...
B: ...

Now, for each item below indicate which is more
    similar to the target domain, in other words,
    with the 20 exemplified sentences. Answer
    sentence A or B, or answer N if it is a tie.
    VERY VERY IMPORTANT: Considering which is more
     similar to the TARGET, in other words, the 20
     exemplified sentences, not the original
    sentence.
a) Topic
b) Style
c) Similar syntactic structure (disregarding
    translation)
d) Named entities
e) Sentiment
f) General meaning
g) Function (informing, narrating, questioning,
    etc.)
h) register


IMPORTANT: Justify. After that, make some comments
     about the adaptation.
```

10.8 Prompt: Translate to a more interpretable prompt

For prompt interpretation, the following prompt was used. In the prompt below, «LEARNED DOMAIN» is exchanged with learned domain embedding.

```
Rewrite the the command below, which is between
    brackets, to be more legible and interpretable.
     Be your best to adapt the embedding
    representing the register, syntax, function,
    topic and target language. You can use many
    words to describe the register, syntax,
    function, topic and target language. The
    command is between brackets. BE VERY CAREFUL
    AND VERY, VERY, VERY DETAILED WITH DOMAIN.

************** START OF THE COMMAND **************
"""[Adapt (register, syntax, function and topic)
    and translate the text using the domain/
    language: <<LEARNED DOMAIN>>]"""
************** END OF THE COMMAND **************
```

10.9 Prompt: Compare original and Adapted sentences in Other Tasks

```
Consider the attributes for a domain below:
```

```
1. Topic: is the subject of a text, as financial,
    medical, computer games, etc. Each topic has a
     vocabulary distribution, and a text can be
    viewed as a topic mixture. So, the topic often
     is not preserved when I adapt the phrase to
    other domain.
2. Genre: is a macro concept that incorporates
    function, register, syntax, and style. For
    instance, an e-mail, a product description,
    and a product advertisement can have different
     genres for the same topics, even if they may
    share the same vocabulary and entity names.
    Automatic genre detection in text can analyze
    the relative frequency of different syntactic
    categories, the use of particular characters
    such as punctuation marks, and sentence length
     distribution.
   2.1. function: texts can be grouped by the
       communicative purpose. In other words, the
       text function can be informing, persuading,
        narrating, instructing, etc.
   2.2. register: conventionalized, functional
       configuration of language tied to certain
       broad social situations. Example: a legal
       discourse has technical vocabulary, a rigid
        structure, and an impersonal tone.
   2.3. syntax: the grammatical structures
       typically used in a genre. Example: short
       sentences and predominance of active voice.
   2.4. style: essentially to do with an
       'individuals use of language.

Here are 20 example sentences from a target domain:

1. ...
...
20. ...

I am analyzing an LLM adaptation activity, which
    translates and adapts a sentence.

The original sentence from the source domain was:
...

Now, consider the translated and adapted sentence
    to the exemplified domain/language, in other
    words, the target domain:

Translated and adapted Sentence:
...


Now, for each item below, indicate how the
    translated and adapted sentence is similar or
    different to the original sentence from the
    source and the target domain (target domain =
    the 20 exemplified sentences). Summarize it in
     a table for each item. Presents only the
    table. Be careful: you must present two
    columns, one comparing the adapted sentence
    with the original sentence and the other
    comparing the adapted sentence with the target
     domain. You must describe why it is similar
    or different for each item below!
```

```
a) Topic
b) Style
c) Similar syntactic structure (disregarding
    translation)
d) Named entities
e) Sentiment
f) General meaning
g) Function (informing, narrating, questioning,
    etc.)
h) register
```