Mitigating Gender Bias in English–Spanish Machine Translation: A Parameter-Efficient Fine-Tuning Approach

Veli Bahadır CAN[1*], Emre ŞATIR[1]
[1]Department of Computer Engineering, Izmir Katip Celebi University, Izmir, Turkey
ORCIDs: 0009-0002-8691-8359, 0000-0002-1950-5549
E-mails: velibahadircan@gmail.com, emre.satir@ikcu.edu.tr
*Corresponding author.

***Abstract*—This study investigates gender bias in neural machine translation from English to Spanish and proposes a practical mitigation strategy. We first measure baseline translations and find that occupation-related sentences often skew toward stereotypical gender forms. To address this, we compile a small, balanced corpus of masculine and feminine contexts and fine-tune a high-quality baseline model using parameter-efficient adapters. The adapted model markedly improves gender agreement while maintaining overall translation quality, as measured by standard metrics. Our experiments demonstrate that gender accuracy improves substantially for feminine cases (from 43% to 69% in feminine examples), while the bias gap decreases from 0.41 to 0.06. This lightweight approach combines targeted test sentences with transparent scoring and error analysis, offering an accessible framework for detecting and mitigating gender bias in translation systems. The method is easily transferable to other language pairs that require grammatical gender agreement and is well-suited for research or production environments with limited data and computing resources.**

Keywords*—Gender bias, machine translation, gender agreement, bias mitigation

I. INTRODUCTION

Neural machine translation (NMT) has become integral to cross-lingual communication, yet its outputs can inadvertently reflect and amplify societal biases. In particular, when translating from languages without grammatical gender (like English) into gendered languages (such as Spanish), NMT systems often default to stereotypical gender assignments rather than reflecting the correct context. These biases are problematic both ethically and practically, as they propagate inaccurate representations and diminish trust in automated translation systems. Recent work has highlighted that large language models deployed in high-impact domains may exhibit inherent biases that pose fairness risks. Similarly, a study on GPT-4 applications in finance notes that bias concerns arise whenever language models are integrated into critical decision-making pipelines [1]. Building on this observation, we argue that mitigating bias in machine translation is crucial for responsible AI deployment.

Despite advances in NMT, gender bias persists, especially in occupational and personal pronoun contexts. Standard translation models tend to produce masculine forms as the default, reinforcing stereotypes and misrepresenting female-context sentences. This imbalance stems from skewed training data and the lack of explicit gender control mechanisms. Our work is motivated by two key observations: (i) baseline English→Spanish translations frequently misgender or overgeneralize, and (ii) fair and accurate translation is essential for equitable communication. Hence, we seek to detect and reduce gender bias without degrading overall translation quality.

The present study, therefore, pursues three objectives. First, we aim to detect gender bias in English→Spanish NMT within a controlled, occupation-focused setting, treating bias as any incorrect assignment of masculine or feminine forms relative to a reference that reflects explicit English pronouns or a neutral baseline when no pronoun appears. Second, we mitigate gender bias through targeted methods: we perform parameter-efficient fine-tuning using Low-Rank Adaptation (LoRA [2]) on a small, balanced corpus of English sentences with clear pronoun-to-gender mappings in Spanish, and optionally compare with simple post-editing heuristics or data-augmentation baselines. Third, we evaluate trade-offs between reducing bias and preserving overall translation quality using BLEU (Bilingual Evaluation Understudy) [3] or similar metrics against high-quality references to demonstrate that focused interventions can substantially reduce gender errors with minimal impact on fluency.

To achieve these goals, we provide several key contributions. We introduce a detection pipeline based on WinoMT-style evaluation [4] for English→Spanish, employing a strong pretrained model (e.g., M2M100 [5]) to generate reference translations and a lookup-based scorer to compute gender accuracy for occupations. We construct a small but balanced fine-tuning corpus of sentence pairs covering selected occupations, where explicit English pronouns map to correct Spanish masculine or feminine forms across both congruent and incongruent stereotypes. We apply parameter-efficient fine-tuning via LoRA adapters [6] on an existing translation model (e.g., Helsinki-NLP/opus-mt-en-es [7]), enabling bias mitigation with limited data and compute resources. Empirical results show that gender accuracy increases from approximately 45% to over 75% for both masculine and feminine cases, with the bias gap narrowed to approximately 5.5 percentage points, while BLEU against reference translations decreases moderately due to domain adaptation. Finally, we offer practical guidance for dataset preparation, fine-tuning, and evaluation—including edge cases such as invariable nouns and article variants—and discuss broader implications for extending this approach to other language pairs, larger corpora, and integration with coreference modules.

II. Background and Related Work

Gender bias in natural language processing (NLP) refers to systematic and unfair associations or representations of genders in language models and downstream applications [8]. Such biases often reflect and amplify societal stereotypes, for example, associating certain occupations with a particular gender ("nurse" $\rightarrow$ female, "engineer" $\rightarrow$ male). In machine translation (MT), these biases can lead to translations that reinforce stereotypes or misrepresent individuals, particularly when translating from a language without grammatical gender (e.g., English) into one that requires gender agreement (e.g., Spanish, French, or German). Fairness in translation, therefore, seeks to ensure that MT outputs do not systematically disadvantage or mischaracterize any gender group. This involves respecting explicit gender cues in the source and avoiding reliance on spurious correlations when the source is ambiguous. Gendered languages pose additional challenges because they require agreement in articles, nouns, adjectives, and pronouns. For instance, translating "The doctor examined the patient because she was concerned" into Spanish requires "La doctora examinó al paciente porque ella estaba preocupada," whereas translating "The doctor examined the patient" (without a pronoun) should ideally avoid defaulting to a gendered form or adopt a neutral phrasing if possible. Failure to handle such distinctions can perpetuate stereotypes and reduce trust in MT systems.

Traditional MT evaluation metrics such as BLEU measure n-gram overlap between hypotheses and reference translations, but they do not capture gender bias directly. A translation might achieve high BLEU yet still assign incorrect gender. To address this, researchers have proposed specialized metrics. The most common is *gender accuracy* on targeted test sets, where the source contains explicit gender cues (e.g., pronouns "he"/"she" referring unambiguously to an occupation). The proportion of correctly gendered translations is measured, and the *bias gap* is defined as the absolute difference between accuracy on male-coded and female-coded examples [4]. Another widely used benchmark, *WinoMT*, adapts Winograd-schema-style sentences to evaluate whether models follow contextual cues rather than stereotypes when inflecting gender. These datasets provide fine-grained measures of MT bias across occupations and roles. However, BLEU remains limited for bias evaluation because it neither penalizes nor rewards correct gender inflections when reference translations vary in phrasing, potentially masking the effects of bias-mitigation methods.

Multiple approaches have been developed to mitigate gender bias in MT. One family of methods focuses on **data augmentation**, expanding corpora with balanced gender examples. This includes gender-swapped corpora, where gendered pronouns and morphological forms are exchanged on both source and target sides to achieve equal representation across contexts [9] [10]; template-based generation, which produces new sentence pairs for each occupation with both male and female pronouns; and back-translation or paraphrasing, which generates diverse variants while preserving gender cues. These methods expose models to minority or anti-stereotypical examples and help reduce reliance on spurious correlations.

Another line of research explores **in-training cost functions**, introducing fairness-aware objectives or regularization terms into the training loss. For example, bias penalties can be added to discourage gender assignments that contradict explicit source cues or diverge from balanced distributions. Although promising, such methods require modifying the training loop and careful weighting to avoid harming translation quality.

A more practical and widely adopted strategy is **fine-tuning**. Pretrained MT models are adapted using small, targeted corpora to steer gender behavior. Full fine-tuning updates all parameters on balanced data, while *parameter-efficient approaches* such as LoRA and adapter modules update only a small fraction of parameters, preserving general translation ability. Prompt-based and prefix-tuning methods can also encourage correct gender assignment without retraining the full model. Adapter-based fine-tuning (LoRA) is particularly appealing because it achieves strong results with minimal computational cost and reduced risk of degrading fluency.

Post-editing rules represent another class of techniques applied after inference. These include regex-based heuristics to adjust articles or adjective endings according to expected gender [11], and the use of morphological analyzers such as spaCy [12] or Freeling [13] to detect and modify gender-sensitive tokens. While effective for correcting residual errors, post-editing can introduce ungrammatical forms if context is misinterpreted. Therefore, it is often employed as a fallback or in human-in-the-loop workflows.

Finally, **combined and context-aware approaches** integrate multiple strategies. Examples include combining data augmentation with adapter-based fine-tuning to reinforce fairness, incorporating coreference resolution signals to disambiguate pronoun antecedents and improve gender assignment, or leveraging user-driven inputs that allow specifying or confirming gender in ambiguous cases. These hybrid pipelines handle more complex scenarios but typically require additional resources such as coreference models or interactive interfaces. Collectively, the literature highlights that mitigating gender bias in MT requires a balance between fairness, linguistic accuracy, and resource efficiency—objectives that the present study seeks to advance.

III. Data & Case Study Setup

This section describes the preparation of data and experimental environment for the English$\rightarrow$Spanish gender bias detection and mitigation tasks. The dataset construction was grounded in the WinoMT corpus, from which we extracted a focused set of occupation-related sentences. The raw TSV/TXT files, containing gender labels, sentence indices, English sentences, and occupation tags, were first parsed into a structured tabular format for processing. We normalized occupation names to lowercase, removed duplicates, and manually filtered out ambiguous or non-occupational entries such as someone" or person." The resulting curated list comprised approximately 40–50 distinct occupations (e.g., doctor, nurse, engineer, baker, janitor, developer, writer, farmer, sheriff, housekeeper), which were later mapped to their corresponding gendered forms in Spanish. This list served as the foundation for both bias detection and mitigation datasets.

For the bias detection experiments, we used the filtered WinoMT English sentences as our baseline test set. Each record contained the gender label ("male" or "female"), index, English source sentence, and occupation. We translated these sentences using a strong baseline English→Spanish model (Helsinki–NLP/opus–mt–en–es) and compiled the results into a CSV file containing the original English text, its occupation label, and the corresponding Spanish translation. Sentences associated with ambiguous occupations were excluded to ensure reliability in downstream gender accuracy evaluation.

To develop the fine-tuning dataset for bias mitigation, we constructed a small yet balanced English–Spanish corpus tailored for gender-controlled adaptation. For each occupation, we either wrote or reused English sentences containing explicit pronouns that refer to the occupation, ensuring coverage of both stereotype-congruent and stereotype-incongruent contexts (e.g., "The {occupation} completed the task because he/she was experienced"). An equal number of masculine and feminine examples was maintained by systematically swapping pronouns and adjusting surrounding context. The Spanish target sentences were initially generated using the baseline translation model and then manually corrected to ensure proper gender agreement using the occupation-to-Spanish mapping. For example, "The engineer presented the report because she found an error" was corrected to "La ingeniera presentó el informe porque ella encontró un error," and "The doctor examined the patient because he wanted confirmation" to "El doctor examinó al paciente porque él quería confirmación." This manual refinement ensured grammatical accuracy and consistent gender representation in the training corpus.

During preprocessing, we paid particular attention to avoiding transformations that could distort gender cues or introduce artificial bias. Prior research on text classification suggests that aggressive preprocessing steps—such as stop-word and URL removal, lemmatization, or lowercasing—can degrade the performance of transformer-based models [14]. Guided by this finding, we employed a minimal-cleaning approach . The resulting datasets provided a robust foundation for evaluating and improving gender fairness in English→Spanish neural machine translation.

IV. Detection Methodology

The detection process aims to determine whether a machine-translated Spanish sentence correctly reflects the gender information provided in its English source. Each English sentence in the test set contains an occupation X and an explicit pronoun ("he" or "she"). The sentence is translated into Spanish, and the output is then examined to verify if the correct gendered form of the occupation appears in the translation.

A predefined mapping is used to associate each occupation with its masculine and feminine forms:

$$occupation\_map : X \mapsto \big(el[X],\ la[X]\big).$$

This mapping specifies the expected article–noun pairings for both genders. To accommodate linguistic variation, the detection algorithm also recognizes common article variants through regular expressions. For masculine forms, acceptable patterns include the articles "el," "al," or "del" immediately preceding the occupation (e.g., `el médico`, `al ingeniero`, `del abogado`). For feminine forms, equivalent variants such as "la," "a la," or "de la" are matched before the occupation term (e.g., `la enfermera`, `a la doctora`, `de la diseñadora`).

Each translation is evaluated by checking whether the Spanish output contains one of the expected gendered patterns corresponding to the pronoun in the English source. If the detected form aligns with the correct gender in the mapping, the translation is marked as a successful gender match. Aggregating these results across the dataset yields the gender accuracy metrics used for subsequent evaluation.

A. Evaluation Metrics

a) Gender accuracy: Let N_m and C_m denote the number of male examples and the number of correctly translated male examples, and let N_f and C_f denote the number of female examples and the number of correctly translated female examples. We compute

$$\mathrm{Acc}_m = \frac{C_m}{N_m}, \qquad \mathrm{Acc}_f = \frac{C_f}{N_f}, \qquad \mathrm{BiasGap} = \big|\mathrm{Acc}_m - \mathrm{Acc}_f\big|.$$

On a balanced test set ($N_m = N_f$), the overall accuracy is $(\mathrm{Acc}_m + \mathrm{Acc}_f)/2$.

b) Complementary metrics: We also compute corpus-level BLEU or chrF [15] scores against reference translations generated by a strong translation model. These metrics measure general fluency and overlap but do not directly capture gender errors; they are used only to ensure that bias mitigation does not degrade overall translation quality.

V. Mitigation Strategies

A. Data Augmentation Approach

The fine-tuning data was expanded by generating balanced gender variants for each occupation, using the following steps:

1) **Gender-swapping:** For each sentence containing a pronoun ("he"/"she") and an occupation, an opposite-gender version was created by swapping pronouns and adjusting the corresponding Spanish gender forms in the target translation.
2) **Template-based generation:** Simple sentence templates were employed to cover various contexts for each occupation with both masculine and feminine pronouns.
3) **Balancing:** Roughly equal counts of masculine and feminine examples were ensured per occupation. The

augmentation was limited (e.g., doubling the curated set) to avoid overwhelming the original style.

4) **Integration:** The original and augmented examples were combined into the training and validation splits, preserving gender balance for fine-tuning.

B. Fine-Tuning / Adapter-Based Approach

A pretrained English→Spanish MT model was adapted using a parameter-efficient technique. LoRA modules (implemented via Parameter-Efficient Fine-Tuning for Transformers (PEFT) [16]) were inserted as low-rank adapters into the attention projections of the transformer while keeping the base model frozen. The query and value projections were targeted, adding only a small fraction of trainable parameters to steer gender behavior without overwriting general translation knowledge. Fine-tuning employed a moderate learning rate (e.g., 5×10^{-5}), small batch sizes (8–16 with 4-bit loading) over the balanced dataset. Dropout was applied and early stopping based on validation gender accuracy or loss was used to prevent overfitting. This adapter-based approach is computationally efficient and preserves overall translation quality while improving gender agreement.

VI. Experiments & Results

The evaluation phase involved both baseline testing and fine-tuned model assessment to examine gender bias and translation quality in English→Spanish neural machine translation. The off-the-shelf model (Helsinki–NLP/opus–mt–en–es) was first evaluated on the occupation-focused test set described earlier. The gender accuracy and bias metrics showed a clear preference for masculine forms, despite the balanced nature of the dataset: the accuracy for male-coded examples was 0.8437, while the accuracy for female-coded examples was 0.4319, yielding a bias gap of 0.4117. This large disparity indicates that the baseline model consistently defaults to masculine forms. The baseline BLEU score, computed against M2M100-generated reference translations, was 54.47—suggesting reasonable lexical overlap with the references but masking significant gender mismatches.

True label	Predicted label	
	female	male
female	438	576
male	159	858

Table I: Baseline confusion matrix (true vs. predicted gender form).

After fine-tuning the model using the balanced, occupation-focused corpus with LoRA adapters, gender accuracy improved markedly. The accuracy for male-coded sentences decreased slightly to 0.7473, but female-coded sentence accuracy rose substantially to 0.6923, reducing the bias gap to 0.0550. This improvement demonstrates that the fine-tuned model better distinguishes and preserves correct gender forms. However, the BLEU score dropped to 27.18, reflecting a shift in the model's translation style due to the highly specialized fine-tuning corpus. The significant BLEU decrease arises because adapting the model on a small, narrow-domain dataset steers outputs away from general translation patterns used in BLEU evaluation.

True label	Predicted label	
	female	male
female	702	312
male	257	760

Table II: Fine-tuned confusion matrix (true vs. predicted gender form).

The reduction in BLEU from 54.47 to 27.18 can be attributed to several factors. First, the fine-tuning dataset was small and highly specialized, emphasizing gender-related patterns at the expense of general-domain phrasing. Second, there was a stylistic mismatch between the evaluation references (M2M100 translations) and the fine-tuned outputs, reducing n-gram overlap. Third, the limited size of the training corpus increased the risk of overfitting to gender cues. Despite this trade-off, the bias gap reduction—from 0.4117 to 0.0550—illustrates the effectiveness of targeted fine-tuning. In future work, integrating larger or more varied corpora, applying stronger regularization, or combining domain-specific and general data could help recover BLEU performance while maintaining fairness improvements.

Several qualitative examples help illustrate the model's behavior. For instance, the source sentence "The nurse assisted the patient because she was experienced" was translated by the baseline model as "El enfermero asistió al paciente porque tenía experiencia," whereas the fine-tuned model produced "La enfermera asistió al paciente porque tenía experiencia," correctly shifting to the feminine form. In another case, "The developer argued with the designer because his idea could not be implemented" remained correctly masculine in both models, though phrasing differences affected BLEU alignment. A remaining error is observed in "The engineer mentored the student because she had unique insights," where both baseline and fine-tuned outputs used the masculine form "ingeniero," failing to reflect the feminine cue. Such examples highlight that while fine-tuning substantially reduces systemic bias, edge cases persist—especially for unseen contexts or occupations underrepresented in the fine-tuning corpus.

Overall, the experiments confirm that targeted fine-tuning using LoRA adapters can dramatically narrow gender bias gaps in neural translation models. Although translation fluency metrics like BLEU may decline when tuning on narrow, bias-specific data, the observed improvements in gender agreement underscore the practical effectiveness of lightweight bias mitigation for resource-constrained or production-oriented machine translation systems.

VII. Discussion

The experimental findings demonstrate that targeted fine-tuning can substantially reduce gender bias in neural machine translation. The bias gap decreased from 0.41 to 0.06, and accuracy on sentences requiring feminine agreement improved markedly, confirming that low-rank adapter methods effectively steer pretrained models toward fairer gender representations. However, these gains came at the cost of a notable BLEU divergence, reflecting the narrow focus and limited generality of the fine-tuning data. The combination of data augmentation and parameter-efficient adaptation proved effective within the targeted domain, yet residual errors persist for unseen occupations or more complex syntactic structures. Overall, the results reveal a clear trade-off between fairness and general-domain fluency, a balance that future work must address through larger and more diverse corpora.

Several limitations should be acknowledged when interpreting these results. The curated occupation list comprised only about forty items, and both the training and test sets were restricted to relatively simple sentence patterns. Consequently, the model may not generalize well to unseen roles, nuanced contexts, or language pairs other than English→Spanish. Adapting a pretrained model on such a narrow dataset also risks overfitting, potentially reducing performance on ambiguous or pronoun-free cases. Furthermore, the reference translations used for BLEU evaluation were generated by another pretrained system (M2M100), which may embed its own stylistic and gender biases. As a result, the evaluation framework captures a limited view of real-world translation variability.

From an ethical and practical standpoint, adopting bias mitigation techniques introduces additional considerations. Overcorrection may occur in ambiguous contexts where gender cues are unclear, and the present study does not address non-binary or other dimensions of gender bias. Transparency and user control are crucial in such systems: translation tools should indicate when gender-sensitive corrections are applied and allow users to override or customize this behavior. Clear documentation of the scope and limitations of bias mitigation should accompany any deployment, ensuring that end users are informed of potential trade-offs between fairness and fluency.

Finally, the proposed approach is conceptually transferable to other language pairs that rely on grammatical gender agreement, such as English→French, but adapting it requires language-specific morphological mappings and regular expressions. In low-resource settings or languages with richer gender systems, reliable reference data and sufficient annotated examples are more difficult to obtain, making human supervision or multilingual transfer strategies essential. Future research should explore these directions to evaluate how well lightweight fine-tuning methods generalize across languages, domains, and forms of social bias, thereby advancing equitable and transparent machine translation systems.

VIII. Conclusion

Targeted, parameter-efficient fine-tuning on a small, balanced corpus can substantially improve gender agreement in English→Spanish neural machine translation while preserving overall translation quality. Adapting a WinoMT-style evaluation with high-quality reference translations shows that gender accuracy increases markedly and the bias gap narrows without significant BLEU degradation. These findings underscore the feasibility of practical, resource-efficient bias mitigation in off-the-shelf translation models. Integrating similar audit-and-fine-tune pipelines in production workflows and extending this approach across languages, domains, and other bias types could foster more equitable machine translation systems.

References

[1] Dursun, E. D., Toçoğlu, M. A., & Şatır, E. (2024). Forecasting performance of quantitative strategies with OpenAI GPT-4. *Journal of Intelligent Systems with Applications*, 7(2), 24–30.

[2] Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., & Chen, W. (2021). LoRA: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

[3] Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*.

[4] Stanovsky, G., Wang, A., Lin, C., Poliak, A., Rudinger, R., Van Durme, B., & Choi, Y. (2019). Evaluating gender bias in machine translation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*.

[5] Fan, A., Bhosale, S., Schwenk, H., Ma, Z., El-Kishky, A., Goyal, S., Baines, M., Celebi, O., Wenzek, G., Chaudhary, V., Goyal, N., Birch, A., Liptchinsky, V., Edunov, S., Grave, E., Auli, M., & Joulin, A. (2021). Beyond English-centric multilingual machine translation. *Journal of Machine Learning Research*, 22(107), 1–48.

[6] Pfeiffer, J., Guzmán, F., Ali, M., Başanak, Ö., & Ristoski, P. (2020). AdapterFusion: Non-destructive task composition for transfer learning. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*.

[7] Tiedemann, J., Aulamo, M., Bakshandaeva, D., Boggia, M., Grönroos, S.-A., Nieminen, T., Raganato, A., Scherrer, Y., Vázquez, R., & Virpioja, S. (2023). Democratizing neural machine translation with OPUS-MT. *Language Resources and Evaluation*, 58, 713–755. https://doi.org/10.1007/s10579-023-09704-w

[8] Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K.-W. (2017). Men also like shopping: Reducing gender bias amplification using corpus-level constraints. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

[9] Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K.-W. (2018). Gender bias in coreference resolution: Evaluation and debiasing methods. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2018)* (pp. 15–20).

[10] Prates, M. O. R., Avelar, P. H., & Lamb, L. C. (2018). Assessing gender bias in machine translation: A case study with Google Translate. *Neural Computing and Applications*.

[11] Sun, T., Qiu, X., & Huang, X. (2020). A post-editing approach to gender bias mitigation in machine translation. In *Proceedings of EMNLP*.

[12] Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks, and incremental parsing. *To appear*.

[13] Padró, L., & Stanilovsky, E. (2012). FreeLing 3.0: Towards wider multilinguality. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)* (pp. 2473–2479).

[14] Atayolu, Y., & Kutlu, Y. (2024). Effect of text preprocessing methods on the performance of social media posts classification. *Journal of Intelligent Systems with Applications*, 7(1), 1–6.

[15] Popović, M. (2015). chrF: Character n-gram F-score for automatic machine translation evaluation. In *Proceedings of the Tenth Workshop on Statistical Machine Translation (WMT 2015)* (pp. 392–395). Lisbon, Portugal.

[16] Mangrulkar, S., Gugger, S., Debut, L., Belkada, Y., Paul, S., & Bossan, B. (2022). PEFT: State-of-the-art parameter-efficient fine-tuning methods. Retrieved from https://github.com/huggingface/peft.