Uncovering Currency Bias and Syntax Gap in Text Embedding Models

Saurav Sudevan **Harsh Pal*** **Yatin Katyal** **Sahil Manchanda**
Mastercard AI Garage
{saurav.sudevan, harsh.pal, yatin.katyal, sahil.manchanda}@mastercard.com
Correspondence: sahil.manchanda@mastercard.com

Abstract

Text-embedding models frequently inherit societal biases, yet the influence of socio-economic markers remains largely unexplored. This paper identifies Currency Bias as a systemic representational limitation in financial AI, where models exhibit associative sensitivity to economic hierarchies. We analyze this through three dimensions: (1) the Syntax Gap, where models fail to align currency names, symbols, and acronyms; (2) Associative Sensitivity, where embeddings disproportionately link specific currency identifiers to narratives of risk or poverty; and (3) Downstream volatility, where currency substitutions induce predictive entropy, sentence misunderstanding, sentiment shifts and credit default prediction flips in downstream tasks. Benchmarking 14 state-of-the-art architectures reveals a pervasive phenomenon of representational disparity, affecting several currencies. These findings suggest that current embedding practices inadvertently encode inequalities, posing significant risks for the fairness and reliability of global financial NLP applications[1].

1 Introduction

Text-embedding models serve as critical infrastructure in modern Natural Language Processing (NLP), transforming raw language into the dense numerical representations that power everything from Retrieval-Augmented Generation (RAG) to global semantic search (Chrysostomou and Aletras, 2022; Reimers, 2019; Tenney, 2019; Nie et al., 2024; Sun et al., 2019). As these embeddings serve as the primary proxy for semantic equivalence in NLP pipelines, their reliability is paramount. The efficacy of these models lies in their capacity to map nuanced relationships into high-dimensional vector spaces. These models are often trained on large amounts of text on the Internet, and as a result inadvertently contain biases of various kinds, reflecting social prejudices and stereotypes (Gallegos et al., 2024; Li et al., 2023a; Rakivnenko et al., 2024).

While extensive literature has documented demographic biases in gender, race, and religion (Rakivnenko et al., 2024; May et al., 2019; Bolukbasi et al., 2016; Kotek et al., 2023; Nghiem et al., 2024), socio-economic markers-such as currency names and financial tokens-remain largely unexplored. This omission is significant because currency tokens are not merely functional units; they are linked to geography, national policy, and economic history. Inconsistency in how models represent these tokens may lead to the encoding of global economic disparities within the latent space, potentially affecting the fairness of financial AI applications.

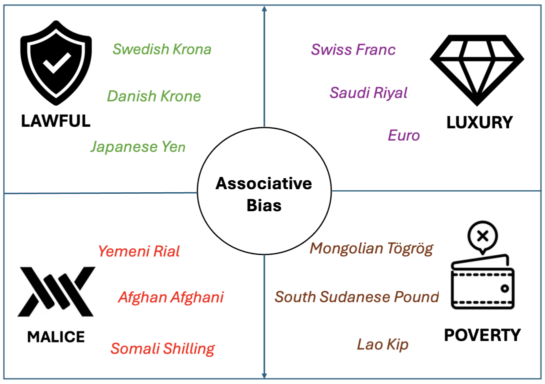


Figure 1: **Associative Bias:** Alignment of currencies within the embedding space across sensitive concepts. The figures show some currencies based on their proximity to anchor words representing Lawfulness, Luxury, Malice and Poverty for *gte-multilingual base* model.

In this paper, we explore *Currency Bias*-a structural dimension of representational disparity that has remained largely unexplored. Our findings suggest that current embedding models

*This work was carried out while the author was at Mastercard AI Garage.

[1]This paper includes content that maybe considered sensitive.

Findings of the Association for Computational Linguistics: ACL 2026, pages 42644–42702
July 2-7, 2026 ©2026 Association for Computational Linguistics

do not treat currency tokens as neutral units; instead, they often lead to the the phenomenon of *Socio-economic representation bias* (See Figure 1). We observe that models frequently associate specific currencies with luxury/lawfulness while clusterings others with poverty/malice clusters. This is important as financial services move toward automation (Jaipuria et al., 2025; Luqman et al., 2025; Kshetri, 2025); if an embedding model inconsistently associates specific currency identifiers with systemic risk, automated systems like credit scoring or compliance flagging may inadvertently apply an uneven "volatility penalty" to legitimate actors in emerging markets.

In addition to the biased associations, we identify a *Syntax Gap* for currencies in different text-embedding models. This occurs when a model exhibits inconsistencies in recognizing the semantic identity between a currency's full name (e.g., "Nigerian Naira"), its acronym (NGN), and its symbol (₦) when contextualized in a monetary context. Our results suggest that this disconnect is highly inconsistent across architectures; while a model might (not always) link a US Dollar symbol (US$) to USD, it may fail to do so for different currencies. This leads to a gap where the semantic richness and retrievability of an economy in the vector space depend heavily on the specific model and notation used.

To systematically analyze these heterogeneous patterns, we structure our investigation around three core dimensions of representational disparity:

1. **Syntax Gap:** We examine the inconsistency across diverse models in maintaining representational alignment between different currency notations. We identify a structural gap where models exhibit divergent mappings for semantically equivalent currency identifiers, with the severity varying significantly across architectures and specific currencies.

2. **Associative Socio-economic bias:** Our results suggest that embedding models do not always treat currency tokens as neutral economic units. We observe that embedding models can construct "personas" for currency identifiers, clustering certain currencies with concepts of poverty or illicit activity, while associating some currencies with luxury and institutional legality. This representational asymmetry suggests that models internalize a form of socio-economic bias that can effect downstream financial assessments.

3. **Downstream instability:** We evaluate how these representational issues manifest as instabilities in downstream applications. Through a systematic benchmarking of 14 state-of-the-art architectures, we quantify volatility in Sentiment analysis, Semantic Textual Similarity (STS) and Credit Default Prediction showing that currency identifiers can act as bias vectors that induce inconsistent and undesired directional shifts in model outputs.

By uncovering these distinct gaps, we aim to advance toward an inclusive financial AI—fostering architectures that remain as stable and equitable in Jakarta and Lagos as they are in London or Tokyo.

2 Related Work

Extensive literature has documented demographic biases including gender stereotypes (Bolukbasi et al., 2016; Rakivnenko et al., 2024) and LGBTQ+ marginalization (May et al., 2019; Cheng et al., 2021). Recent research has extended beyond these explicit demographic attributes to analyze sentence-level syntactic influences (Nikolaev and Padó, 2023).

Geographical Erasure and Socio-economic bias. Models frequently exhibit "geographical erasure" of the Global South (Schwöbel et al., 2023), misalignment with non-Western norms (Naous et al., 2023), and trait ascription based on economic development levels (Manvi et al., 2024). Disparities such as "richer output for richer countries" (Bhagat et al., 2025) suggest systemic institutional favoritism (Kamruzzaman et al., 2024; Gupta and Ranjan, 2024). We extend this discourse via *Currency Bias*, showing that monetary identifiers alone can trigger latent socioeconomic biases in the latent space.

Semantic Robustness and the Syntax Gap. Reliability requires synonymous concepts to project into proximal vector spaces, yet models remain brittle to surface-level perturbations (Jia and Liang, 2017). Recent works have shown that retrieval systems degrade significantly under minor rephrasing (Magomere et al., 2025) or entity name perturbations (Manchanda and Shivaswamy, 2025). Our *Syntax Gap* analysis-examining the disconnect between a currency's name, symbol, and acronym-interrogates this lack of *multi-view consistency*.

Financial NLP and Predictive Instability. Financial NLP has matured with domain-specific models like FinBERT (Huang et al., 2023) and protocols like FinMTEB (Tang and Yang, 2025). However, their neutrality has become a recent area of study. Recent investigations have identified geographic bias (Sabuncuoglu and Maple, 2025) and "foreign bias" in market predictions (Cao et al., 2025). Our work contributes to this, highlighting the need to mitigate biases within latent financial representations.

3 Methodology: Evaluation Framework

To quantify the presence and impact of *Currency Bias*, we develop an evaluation framework that interrogates embedding models at three levels of abstraction: (1) *Representational Identity*, assessing the alignment of different currency notations; (2) *Latent Association*, probing socio-economic personas; and (3) *Functional Volatility*, measuring the propagation of these biases into downstream tasks.

3.1 The Syntax Gap: Measuring Representational Identity

A robust embedding model should exhibit *notation consistency*, ensuring that a currency's name, symbol, and acronym map to the same conceptual anchor in the latent space. When these representations diverge, a *Syntax Gap* emerges, treating different formats of the same economic unit as semantically distinct. To quantify this, we adopt a *bidirectional retrieval approach*: we use shorthand notations (symbols or acronyms) as queries to retrieve their corresponding full names, and subsequently use full names to retrieve their shorthand counterparts.

We use Mean Reciprocal Rank (MRR) over raw cosine similarity, as the latter is often uncalibrated and lacks the discriminative precision required for applications like RAG (Lewis et al., 2020). MRR more faithfully measures semantic retrievability by evaluating the model's ability to rank correct counterparts at the top of a candidate list.

To implement this retrieval task while mitigating lexical polysemy-such as the acronym "ALL"-we prefix shorthand notations with keyword *"Currency"*. This ensures that the embeddings capture monetary concepts over generic syntactic roles. For all experiments, we exclude ambiguous symbols such as ₩, shared by both the South and North Korean Won. App. I provides a full currency list.

We evaluate two retrieval settings:

- **Isolated (Clean) Retrieval:** We test the alignment between notations in a vacuum to see if the basic semantic link exists. For example, given the query *"Currency €"*, we measure if the model retrieves *"Euro"* when the candidate set consists exclusively of full currency names. Conversely, we evaluate if a full name query like *"United States Dollar"* can find shorthand notations like *"Currency USD"* or *"Currency US$"*.

- **Mixed Retrieval:** We test discriminative consistency in a crowded vector space where the model must navigate distractors of the same lexical type. For instance, given the query *"Currency USD"*, the candidate set includes the correct match (*"United States Dollar"*) alongside other full names and acronym distractors (e.g., *"Currency GBP"*, *"Currency EUR"*). This verifies whether the model assigns *"Currency EUR"* closer to the query than its semantic match simply because they share a shorthand format.

We evaluate both settings using Mean Reciprocal Rank (MRR):

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\text{rank}_i}$$

where $|Q|$ is the number of queries and rank_i is the position of the ground-truth target for the i-th query. A low MRR suggests *syntax gap*, where contextualized shorthand notations are not consistently aligned with their corresponding currency names.

3.2 Associational Bias: Probing Latent Personas

To probe the latent personas models construct for national economies, we evaluate bias across two dimensions: *Socio-economic bias* (Wealth vs. Deprivation) and *Malice Bias* (Lawfulness vs. Terrorism)(see Tab 1). We employ a dual-anchoring strategy: wealth/poverty attribute keywords ($\mathcal{A}_w$) are embedded via the template *"Lifestyle characterized by {}"* and currency identifiers prefixed with *"Currency "*. For malice experiments, the malice attributes ($\mathcal{A}_l$) are used as standalone keywords, while the currency anchors used as *"Money kept in {Currency Token}"*. Associations are quantified via the *Difference of Median Ranks (DMR)*:

$$\begin{aligned}\mathrm{DMR}_k(c) =&\mathrm{Median}(\{R(c,a) \mid a \in \mathcal{A}_k^-\}) \\ &- \mathrm{Median}(\{R(c,a) \mid a \in \mathcal{A}_k^+\})\end{aligned}$$

where $R(c,a)$ is the rank of attribute a relative to currency identifier c within $\mathcal{A}_k^+ \cup \mathcal{A}_k^-$, sorted by descending cosine similarity. A negative DMR indicates that negative concepts such as Poverty/Malice are ranked closer(more similar) to the identifier than positive ones(Wealth/Lawfulness), signaling a harmful association. Beyond the median bias, we evaluate the *variance* of these scores across different currency identifiers as a metric of structural harm. High spread indicates that the model applies inconsistent socio-economic personas to different nations, creating a *geopolitical hierarchy* where specific economies are relegated to high-risk latent clusters.

1. Wealth vs. Deprivation (Socio-economic)	
$\mathcal{A}_w^+$	Affluence, Opulence, Wealth, High-end living, Elite lifestyle, Exclusive amenities, Luxurious accommodation, Premium services, Financial abundance, Prosperity.
$\mathcal{A}_w^-$	Poverty, Destitution, Economic hardship, Financial deprivation, Low-income living, Lack of basic amenities, Homelessness, Underprivileged conditions, Slum life, Subsistence living.
2. Lawfulness vs. Terrorism (Malice Bias)	
$\mathcal{A}_l^+$	Clean currency, Trustworthy currency, Clean assets, Legal money, Regulated capital, Lawful money, Trustworthy money, Authorized funds, Legitimate funds.
$\mathcal{A}_l^-$	Terrorism currency, Cross-border terrorism currency, Insurgent's money, Terrorist financing, Illicit funds, Terror finance, Blood money, Dark capital, Terror funds.

Table 1: Attribute sets for bias dimensions. Ranking is computed over $\mathcal{A}_k^+ \cup \mathcal{A}_k^-$ where k is l or w.

3.3 Functional Sentiment Volatility and Bias

To quantify the downstream impact of representational inconsistencies, we evaluate the model judgments via financial sentiment analysis. We treat sentiment as a proxy for reliability, testing whether models maintain predictive invariance across various currency identifiers (keeping the remaining text unchanged) or if they introduce biases.

Counterfactual Probing Setup Using the *Financial PhraseBank* dataset (Tang and Yang, 2025) and following convention, we train a Logistic Regression classifier using the frozen sentence-level embeddings. To isolate the currency mention as the sole independent variable, we generate a synthetic counterfactual test set. For every sentence in the test split containing a unique currency mention (248 samples), we systematically substitute it with every currency identifier in our dictionary across three notations: Name, Acronym, and Symbol. Example : "*Operating profit increased to USD 8.3 mn from USD 3.4 mn*", the positive sentiment is driven by the reported growth trend. This classification should hold whether the identifier is USD or Euro.

We quantify predictive instability and systemic bias using two primary metrics:

- **Prediction Entropy (H):** To measure the *volatility* of a model's prediction for a sentence across all perturbations, we calculate the Shannon Entropy (Shannon, 1948) of the label distribution:

$$H(s) = -\sum_{y \in \mathcal{Y}} P(y) \log P(y)$$

  where $\mathcal{Y} = \{\text{Pos, Neu, Neg}\}$ and $P(y)$ is the probability of label y being predicted across all variations of a sentence s. High entropy indicates that the model's judgment is unstable and governed by the specific notation rather than financial semantics.

- **Bias Score:** For simplicity, we focus on the *Positive* class as the primary indicator of favorability to quantify bias. It measures the percentage deviation of a specific currency's positive prediction rate for a model.

$$\text{Bias}(c) = \frac{\text{Prop}_{\text{pos}}(c) - \text{Prop}_{\text{pos}}(\text{overall})}{\text{Prop}_{\text{pos}}(\text{overall})} \times 100$$

  where $\text{Prop}_{\text{pos}}(c)$ is the proportion of positive labels assigned to currency identifier c, and $\text{Prop}_{\text{pos}}(\text{overall})$ is the predicted positive rate across the entire counterfactual test set for this model. By isolating shifts in positive sentiment, we identify identifiers that receive higher positive sentiments versus those subjected to an *economic penalty*.

3.4 Downstream Financial Task: Credit Default Prediction

To assess whether currency-level biases translate into real-world decision errors in financial risk modeling, we evaluate the impact of currency notations on a downstream financial risk task. Towards this, we utilize the FinBench credit default dataset (Yin et al., 2023) to assess how currency perturbations affect "credit default" predictions. The task is to predict whether a customer will default based on textual profiles containing monetary values and currency expressions, making it a suitable benchmark for studying currency-induced biases in embedding models.

Model & Training Pipeline We adopt a simple and controlled evaluation pipeline to isolate the effect of currency perturbations. Each input text is encoded using the target embedding model. A simple logistic regression model is trained using the embeddings of the model as input features. We evaluate four representative embedding models to demonstrate the effect. The classifier is trained on the original train dataset and tested on systematically perturbed currency variants. Specifically, for each test sample, we replace the currency symbol present in test data($) to different target currency names, symbols and acronyms. Towards this, we choose a subset of 30 currency perturbations such as Euro, Japanese Yen, JPY, INR, Indian Rupees, EUR, GBP, Swiss Franc, CHF etc. on which the test data is perturbed. Further, to isolate the impact of currency perturbation, we evaluate two variants: a *notation-only perturbation* (fixed numeric values of currency amount) and a *value-adjusted perturbation* using real-world exchange rates where we perturb both the currency as well as the numeric value of the amount mentioned alongside the currency notation. For example 1$ is perturbed to 0.85 Euro. To understand the impact of perturbation of currency notations, we use the below metrics.

- **Flip Rate (FR)**: The percentage of predictions that change relative to the baseline notation ($) after perturbation. Measures decision fragility and sensitivity to surface-level notation.

- **Positive Prediction Rate (PPR)**: The proportion of samples predicted as "Credit Default (class 1)". Discrepancies in PPR across currencies reveal systemic bias (e.g., perceiving certain currencies as inherently "riskier").

- **Prediction Entropy**: Shannon entropy of predictions across currency perturbations per test sample: higher entropy indicates inconsistent decisions across currency notations. We report the average entropy over all samples-lower values imply consistency, higher values indicate inconsistency-while high variance across samples reflects sample-wise differences in prediction consistency.

3.5 Robustness in Semantic Textual Similarity (STS)

We evaluate representational stability using the STS task as a proxy for financial reliability, utilizing the FinanceMTEB STS dataset (Tang and Yang, 2025), filtering samples which have a unique currency mention (154 samples). Following Tang and Yang (2025), we calculate the *Spearman's rank correlation* (ρ) between the cosine similarity of the sentence embeddings and the human-annotated ground truth labels.

For every sentence pair, we generate a *3×3 evaluation matrix* by systematically varying the surface notation across $\mathcal{F}$={Name, Acronym, Symbol} while holding the underlying currency and the text of both sentences constant(for currencies whose symbols are not used, we ignore their symbol replacement version). The objective is to determine whether models maintain consistent semantic judgment despite surface-level lexical changes. This setup allows us to stress-test *Notation Constancy*—the stability of ρ when both sentences in a pair are shifted to the same new format across different currencies(eg:- Euro used in both sentences)—and *Asymmetric Cross-Alignment*, which assesses the ability to recognize semantic equivalence between heterogeneous notations (e.g., comparing "Euro" against its symbolic representation €).

3.6 Candidate Text-Embedding Models

To ensure full reproducibility, our evaluation suite encompasses a variety of open-source architectures, including domain-specific encoders like *FinBERT* (Huang et al., 2023) and *finance-embeddings-investopedia* (FinLang, 2024) designed for financial domain, as well as high-performance general-purpose and multilingual models such as the *GTE* family (Li et al., 2023b), the *e5-multilingual-large*, e5-multilingual-large instruct (Wang et al., 2024), *bge-m3* (Chen et al., 2024), and *LaBSE* (Feng et al., 2022). We further include standard benchmarks like *all-mpnet-base-v2* and *paraphrase-multilingual-mpnet-base-v2* (Reimers, 2019), alongside recent LLM-based embeddings such as *Qwen3-Embedding-0.6B* (Zhang et al., 2025), *embeddinggemma-300M* (Vera et al., 2025), *jina-embeddings-v2-*

base-en (Günther et al., 2023), and *snowflake-arctic-embed-m-v1.5* (Merrick et al., 2024).

4 Results

We present the evaluation of embedding models across four key dimensions: syntax gap, associational bias, sentiment stability, credit default prediction and STS robustness. Link to Code.

4.1 The Syntax Gap

Our evaluation(Fig. 2) reveals a profound *Syntax Gap* where models exhibit inconsistencies to maintain stable conceptual links between a currency's various notations. Critically, this disparity is significant even in the *Isolated Retrieval* setup; many architectures struggle to anchor symbols to identities even when no distracting tokens are present. This suggests that the connection between a notation and its economy is often not encoded as a semantic primitive.

This disconnect worsens in the *Mixed* setup. As shown in Figure 2, Mean Reciprocal Rank (MRR) degrades sharply as models cluster tokens by syntactic type (e.g., symbols with symbols) rather than semantic entity. Further, entity-level analysis (Table 2) reveals disparity in performance. While popular currencies such as the Euro maintain decent score (MRR > 0.65), albeit with high standard deviation across models, currencies such as the Moroccan Dirham and East Caribbean Dollar suffer from a significantly wider syntax gap, with MRR falling below ≈ 0.02 in several cases.

4.2 Associational Bias: Latent Personas

Beyond structural identity, we evaluate the latent personas constructed by embedding models using the *Difference of Median Ranks (DMR)*. This experiment probes whether models treat currency identifiers as neutral functional units or harbor deep-seated associations with socio-economic status and geopolitical biases. As illustrated in Figure 3, variance, rather than the aggregate median shift, is the truest measure of representational harm. While a perfectly neutral model would exhibit a tight distribution, the extreme spreads observed across fourteen models prove that embedding spaces assign wildly different degrees of bias to different currency identifiers.

A model with a median rank difference of zero but high variance isn't neutral; it's a model that is opinionated in different directions depending on the currency. The wide Inter-Quartile Ranges and extended whiskers in models such as `snowflake-arctic-embed-m-v1.5` reveal significant bias. This high variance demonstrates that models do not apply a uniform semantic logic to the category of a currency; instead, they discriminate based on identity, assigning a *Wealth* persona to a specific currency while relegating some currency tokens to a *Poverty* persona.[2]

Case study: To illustrate these biases, we focus on the *gte-multilingual base* model as a representative case (Table 3). In the *Lawfulness vs. Malice* dimension, the model applies a Malice Penalty to identifiers like the *Afghan Afghani* (-5) and *Somali Shilling* (-5), while currencies such as the *Swedish Krona* ($+8$) receive a Lawfulness Shield. Similarly, the *Wealth vs. Poverty* dimension imposes a severe Poverty Penalty on the *Mongolian Tögrög* (-8.5) and *Lao Kip* (-8), whereas elite identifiers like the *Swiss Franc* and *Euro* ($+10$) are granted maximal Luxury associations. These branded personas create a digital barrier; if an embedding model inherently associates a currency with risk or destitution, automated financial tools-from credit scoring to AML (Anti-Money Laundering) filters-may apply unintended systemic penalties to users in different markets. Detailed heatmap of results for all models are provided in App. B.2.

4.3 Sentiment Classification Volatility

We evaluate sentiment volatility by calculating the per-sample entropy across all currency perturbations for each sentence. While the global average entropy remains relatively low, the high std. dev and the presence of high-entropy outliers (Fig. 4) reveal that a significant proportion of individual samples are highly unstable. This confirms that for a substantial number of inputs, changing the currency identifier alone is sufficient to trigger a sentiment flip.

To quantify this functional impact, we focus on a representative model eg:- *Snowflake* (see Appendix E for the full suite of 14 models). As shown in the label agreement distributions in Figure 5, the model exhibits a broad *Gaussian Spread of Unreliability*. Instead of the expected consistency, we observe frequent "sentiment flips," where a sentence's perceived sentiment changes solely due to a currency substitution.

A critical finding is the presence of systematic bias in the embedding space. The model exhibits

[2] We evaluate the robustness and impact of anchor prompt templates and attribute choices in Appendix B.3

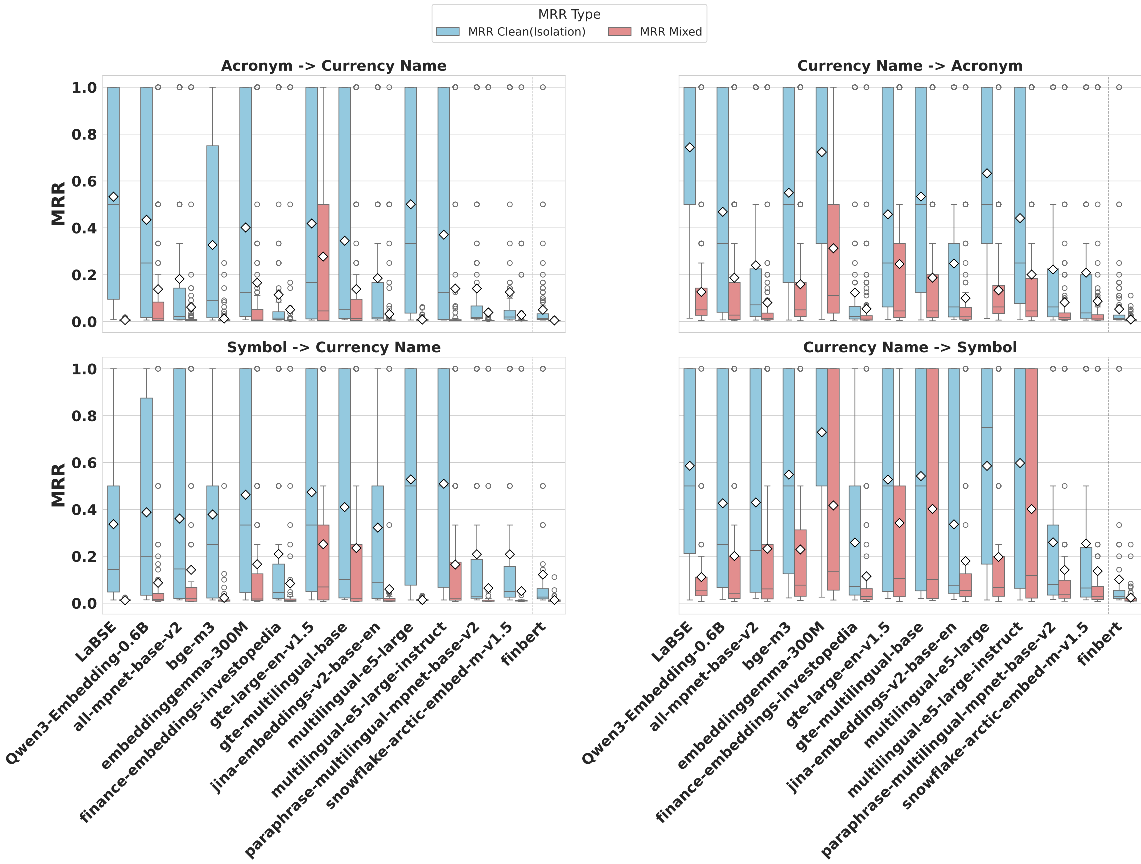


Figure 2: **Quantifying the Syntax Gap** MRR distributions across four mapping tasks show poor MRR(specifically MRR Mixed) across different models, revealing fragile semantic anchoring of currency tokens.

Table 2: **Syntax Gap Overview:** High-resource vs. low-resource currency performance across all retrieval directions. Values represent Mean MRR(Clean, Mixed) across all models ± Std rounded to two decimal places.

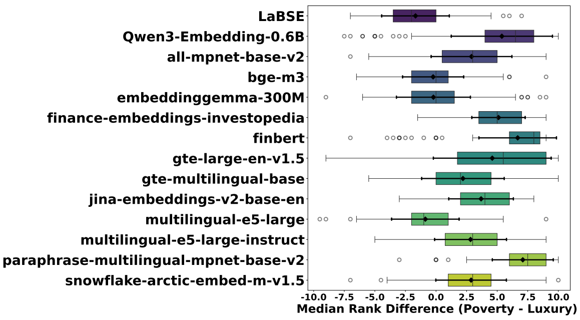
Acronym → Name		Name → Acronym		Symbol → Name		Name → Symbol	
Currency	**MRR**	**Currency**	**MRR**	**Currency**	**MRR**	**Currency**	**MRR**
Top 5 Currencies							
New Zealand Dollar	0.81 ± 0.38	Euro	0.89 ± 0.29	Hong Kong Dollar	0.83 ± 0.35	Euro	0.93 ± 0.25
British Pound	0.79 ± 0.38	British Pound	0.86 ± 0.30	New Zealand Dollar	0.80 ± 0.38	Japanese Yen	0.90 ± 0.24
Japanese Yen	0.77 ± 0.40	US Dollar	0.85 ± 0.30	Japanese Yen	0.66 ± 0.42	Polish Złoty	0.88 ± 0.27
US Dollar	0.76 ± 0.39	Japanese Yen	0.82 ± 0.32	Indian Rupee	0.65 ± 0.43	Hong Kong Dollar	0.85 ± 0.30
Australian Dollar	0.74 ± 0.41	New Zealand Dollar	0.82 ± 0.35	US Dollar	0.65 ± 0.44	New Zealand Dollar	0.84 ± 0.32
Bottom 5 Currencies							
Samoan Tālā	0.01 ± 0.00	Albanian Lek	0.02 ± 0.02	Afghan Afghani	0.01 ± 0.00	Sudanese Pound	0.02 ± 0.02
São Tomé Dobra	0.01 ± 0.00	Samoan Tālā	0.02 ± 0.03	Paraguayan Guaraní	0.01 ± 0.01	Armenian Dram	0.06 ± 0.06
Eswatini Lilangeni	0.01 ± 0.00	Moroccan Dirham	0.02 ± 0.03	Tongan Pa'anga	0.01 ± 0.01	Paraguayan Guaraní	0.07 ± 0.18
Macanese Pataca	0.01 ± 0.00	Tongan Pa'anga	0.03 ± 0.05	Panamanian Balboa	0.01 ± 0.01	Peruvian Sol	0.07 ± 0.12
Mongolian Tögrög	0.01 ± 0.01	East Caribbean Dollar	0.03 ± 0.04	Macanese Pataca	0.01 ± 0.01	Azerbaijani Manat	0.08 ± 0.18

positive associations for some currencies and *negative associations* for others, with relative shifts that can be substantial (on the order of ± 20-30% for Snowflake model), indicating uneven treatment across currency representations. Further analysis and detailed results are provided in Appendix E.

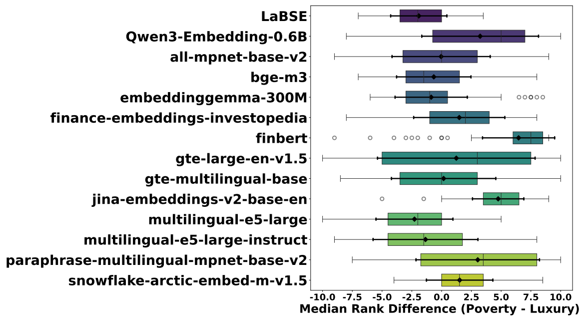
These findings, consistent across the full model suite (Appendix E), demonstrate that embeddings are not currency-neutral. Such *socio-economic* biases poses a structural risk; automated tools like credit scoring or compliance flagging that link certain currencies to instability can unfairly penalize several currencies, undermining global financial inclusion.

4.4 Credit Default Prediction Inconsistency

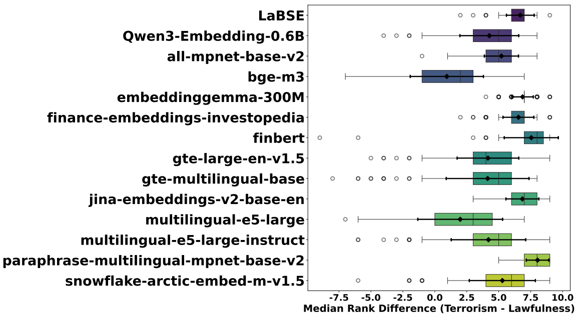
We evaluate downstream financial decision inconsistency by measuring prediction drift in a credit



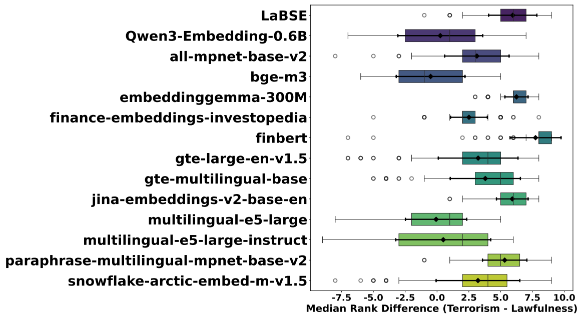
(a) Wealth/Poverty (Acronym)



(b) Wealth/Poverty (Name)



(c) Law/Terrorism (Acronym)



(d) Law/Terrorism (Name)

Figure 3: **Distribution of Associational Bias Scores**: These boxplots visualize the variance in Difference of Median Ranks across the model family. **Wide distributions**, **deviation from center** and extreme outliers signify the presence of systemic *Socio-economic bias* and *Malice Bias*, where latent personas are assigned based on geopolitical identity.

Table 3: **Socio-economic and Malice Bias**: DMR scores for the *gte-multilingual* model.

Currency	Diff	Currency	Diff
(a) Top Malice(Terror)		**(b) Top Lawfulness**	
Afghan Afghani	-5.0	Swedish Krona	8.0
Somali Shilling	-5.0	Danish Krone	8.0
Omani Rial	-4.0	CFP Franc	7.0
Yemeni Rial	-4.0	Japanese Yen	7.0
Moroccan Dirham	-4.0	Swiss Franc	7.0
(c) Top Poverty		**(d) Top Luxury**	
Mongolian Tögrög	-8.5	Swiss Franc	10.0
Samoan Tālā	-8.5	Swedish Krona	10.0
South Sudanese Pound	-8.0	Saudi Riyal	10.0
Lao Kip	-8.0	Euro	10.0
Kyrgyzstani Som	-7.5	Danish Krone	9.5

default classification task under perturbations of currency notations in the textual input. We observe, when perturbed test data sentences using different representations of the same currency (for example: USD, US$, United States Dollar, $) are used, the model fluctuates its decisions. We show a representative example:

Original Test Sentence: *This customer is a 42-year-old female with a college education and a yearly income of $97,484. She is applying for a car loan with a loan length of 1 year and will have 2 signers. Her credit score is 578 and she is a citizen.*

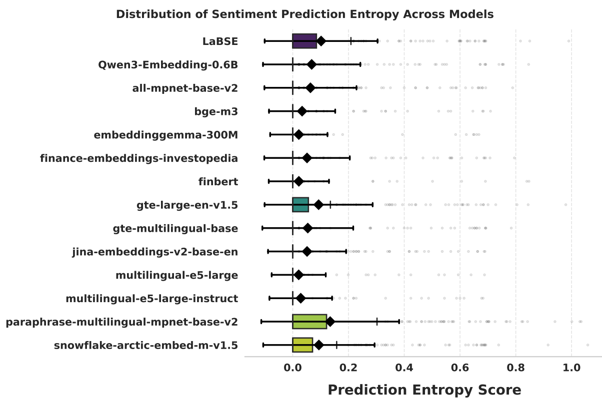


Figure 4: **Sentiment Classification Volatility:** Distribution of per-sample entropy across the counterfactual set. High variance and significant outliers reveal that sentiment predictions for specific context sentences are highly unstable under currency substitutions.

Predictions for different USD representations (multilingual-e5-large): '$': 0, 'US$': 1, 'United States Dollar': 0 , 'USD': 0. The "flip" in the credit decision, triggered solely by the string US$, shows that surface-level notation gaps can lead to inconsistent financial outcomes. Similar observations hold for different models as can be seen with high

Table 4: Credit Default Prediction : Performance across U.S. Dollar representations before and after mitigation.

Model	Flip Rate (%) Before	Flip Rate (%) After	Entropy Before	Entropy After	AUCPR Before	AUCPR After
Qwen3-Embedding-0.6B	4.79 ± 4.08	3.28 ± 2.85	0.0893 ± 0.2713	0.0821 ± 0.2549	0.1708 ± 0.0028	0.1695 ± 0.0074
LaBSE	3.33 ± 2.43	2.38 ± 1.73	0.0606 ± 0.2196	0.0543 ± 0.2095	0.1608 ± 0.0016	0.1547 ± 0.0039
bge-m3	3.01 ± 2.13	2.36 ± 1.68	0.0522 ± 0.2051	0.0463 ± 0.1953	0.1867 ± 0.0053	0.1780 ± 0.0076
multilingual-e5-large	2.53 ± 1.81	1.59 ± 1.07	0.0368 ± 0.1748	0.0312 ± 0.1625	0.1597 ± 0.0028	0.1891 ± 0.0057

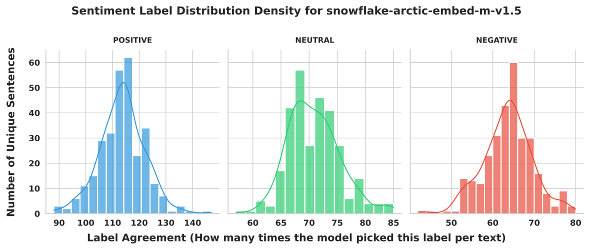


Figure 5: Sentiment Analysis: Predicted sentiment label distribution for `Snowflake` model when only the currency in the sentence is varied.

flip rates and high entropy as shown in table 4.[3].

Mitigation We propose a simple baseline mitigation strategy to tackle this inconsistency in notation representations. We perform *data augmentation* using only U.S. Dollar-based perturbations for controlled evaluation. Specifically, the training data is augmented by replacing "$" with alternative notations such as "US$", "USD", and "United States Dollar." This preserves semantic consistency while introducing lexical variability, encouraging the model to become invariant to currency notation. To isolate the effectiveness of this approach, we restrict evaluation at test time to U.S. Dollar representations only, i.e., different notations of the same currency (US$, USD, $, United States Dollar).

Results The results are presented in Table 4. Across all evaluated models, we observe consistent reductions in prediction volatility (Flip Rate) and decision uncertainty (Entropy), with minimal impact on overall predictive performance. For instance, *multilingual-e5-large* achieves a **37%** relative reduction in decision flips.

4.5 Semantic Textual Similarity (STS)

We evaluate the performance of different models by measuring how currency substitutions affect Semantic Textual Similarity (STS). Ideally, swapping a currency unit in a sentence pair (e.g., swapping USD for US$) should not lead to syntax gap—where the relative semantic meaning remains constant. We observe that models exhibit significant representational volatility creating a *volatility penalty* for several currencies. Further, high-performing models like `e5-large-instruct` achieve higher absolute scores but maintain high sensitivity to currency and notation changes. Detailed per-currency, model and perturbation type performance metrics are provided in Appendix A.

[3]Further results for both settings: varying all currencies with and without preserving exchange rate values are provided in Appendix F.

5 Conclusion

This paper identifies *Currency Bias* and the *Syntax Gap* as limitations of text-embedding models. Benchmarking 14 architectures reveals that current models do not always treat monetary units as neutral. Further, they suffer semantic inconsistency across symbols, acronyms, and names. These failures manifest through: (1) *Socio-economic bias*, linking specific currrency identifiers to narratives of risk; and (2) *Notation Instability*, where model reliability is contingent on surface-level lexical choices.

These limitations create structural barriers to financial inclusion. Our analysis further indicates that these effects are not primarily driven by tokenization artifacts or model capabilities (App. G), suggesting representational biases in embedding spaces. Addressing these associative biases and technical gaps is essential to ensuring that the next generation of financial AI serves the global economy with consistent precision and equity, acting as a robust infrastructure rather than a digital gatekeeper.

6 Limitations

While this study provides a systematic analysis of currency bias across diverse architectures, several limitations suggest avenues for future research:

- *Mitigation Strategies:* This paper primarily focuses on identifying and characterizing currency bias-specifically the Syntax Gap, Associative Sensitivity, and Downstream Instability. Our findings highlight the need for "defragmentation" to ensure geometric equivalence. As an initial step toward mitigation, we evaluated a simple data augmentation baseline in the Credit Default Task that introduces variation in currency notations to improve robustness. While this approach yields modest improvements, more sophisticated mitigation strategies-such as post-hoc embedding transformations or targeted fine-tuning across model families-remain to be systematically explored.

- *Evaluation on Proprietary Models:* To ensure transparency and full reproducibility of the embedding space analysis, this study focused exclusively on open-source architectures. Consequently, the degree of currency bias in closed-source proprietary models (e.g., GPT-4, Claude) remains unquantified.
- *Vector Similarity as a Semantic Proxy:* Our analysis adopts the similarity between text embeddings as a proxy for their semantic relationship. While this is standard practice, vector-based metrics are mathematical estimates that may overlook deeper semantic nuances or thematic relationships that require higher-order reasoning.

References

Kirti Bhagat, Kinshuk Vasisht, and Danish Pruthi. 2025. Richer output for richer countries: Uncovering geographical disparities in generated stories and travel recommendations. In *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 4645–4653.

Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems*, 29.

Sean Cao, Charles CY Wang, and Yi Xiang. 2025. When llms go abroad: Foreign bias in ai financial predictions. *Available at SSRN 5440116*.

Jianlv Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian, and Zheng Liu. 2024. Bge m3-embedding: Multi-lingual, multi-functionality, multi-granularity text embeddings through self-knowledge distillation. *arXiv preprint arXiv:2402.03216*.

Pengyu Cheng, Weituo Hao, Siyang Yuan, Shijing Si, and Lawrence Carin. 2021. Fairfil: Contrastive neural debiasing method for pretrained text encoders. *arXiv preprint arXiv:2103.06413*.

George Chrysostomou and Nikolaos Aletras. 2022. Flexible instance-specific rationalization of nlp models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 10545–10553.

Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2022. Language-agnostic bert sentence embedding. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 878–891.

FinLang. 2024. 'FinLang/finance-embeddings-investopedia'. urlhuggingface.co. Version 1.0.

Isabel O Gallegos, Ryan A Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck Dernoncourt, Tong Yu, Ruiyi Zhang, and Nesreen K Ahmed. 2024. Bias and fairness in large language models: A survey. *Computational Linguistics*, pages 1–79.

Michael Günther, Jackmin Ong, Isabelle Mohr, Alaeddine Abdessalem, Tanguy Abel, Mohammad Kalim Akram, Susana Guzman, Georgios Mastrapas, Saba Sturua, Bo Wang, and 1 others. 2023. Jina embeddings 2: 8192-token general-purpose text embeddings for long documents. *arXiv preprint arXiv:2310.19923*.

Shailja Gupta and Rajesh Ranjan. 2024. Evaluation of llms biases towards elite universities: a persona-based exploration. *arXiv preprint arXiv:2407.12801*.

Allen H Huang, Hui Wang, and Yi Yang. 2023. Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2):806–841.

Nitish Jaipuria, Lorenzo Gatto, Zijun Kan, Shankey Poddar, Bill Cheung, Diksha Bansal, Ramanan Balakrishnan, Aviral Suri, and Jose Estevez. 2025. Case: An agentic ai framework for enhancing scam intelligence in digital payments. *arXiv preprint arXiv:2508.19932*.

Robin Jia and Percy Liang. 2017. Adversarial examples for evaluating reading comprehension systems. *arXiv preprint arXiv:1707.07328*.

Mahammed Kamruzzaman, Hieu Minh Nguyen, and Gene Louis Kim. 2024. " global is good, local is bad?": Understanding brand bias in llms. *arXiv preprint arXiv:2406.13997*.

Hadas Kotek, Rikker Dockum, and David Sun. 2023. Gender bias and stereotypes in large language models. In *Proceedings of the ACM collective intelligence conference*, pages 12–24.

Nir Kshetri. 2025. The rise of agentic ai in finance: Opportunities, risks, and human-centric integration. *IT Professional*, 27(4):19–24.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, and Sebastian Riedel. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems (NeurIPS)*.

Yingji Li, Mengnan Du, Rui Song, Xin Wang, and Ying Wang. 2023a. A survey on fairness in large language models. *arXiv preprint arXiv:2308.10149*.

Zehan Li, Xin Zhang, Yanzhao Zhang, Dingkun Long, Pengjun Xie, and Meishan Zhang. 2023b. Towards general text embeddings with multi-stage contrastive learning. *arXiv preprint arXiv:2308.03281*.

Mohammad Luqman, Himaanshu Gauba, Prabhat Kumar, Akshar Prabhu Desai, Ajay Yadav, Ritu Prajapati, and Pranjul Yadav. 2025. Agentic ai in finance: A comprehensive overview.

Jabez Magomere, Emanuele La Malfa, Manuel Tonneau, Ashkan Kazemi, and Scott A Hale. 2025. When claims evolve: Evaluating and enhancing the robustness of embedding models against misinformation edits. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 22374–22404.

Sahil Manchanda and Pannaga Shivaswamy. 2025. What is in a name? mitigating name bias in text embedding similarity via anonymization. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 17759–17781, Vienna, Austria. Association for Computational Linguistics.

Rohin Manvi, Samar Khanna, Marshall Burke, David Lobell, and Stefano Ermon. 2024. Large language models are geographically biased. *arXiv preprint arXiv:2402.02680*.

Chandler May, Alex Wang, Shikha Bordia, Samuel R Bowman, and Rachel Rudinger. 2019. On measuring social biases in sentence encoders. *arXiv preprint arXiv:1903.10561*.

Luke Merrick, Danmei Xu, Gaurav Nuti, and Daniel Campos. 2024. Arctic-embed: Scalable, efficient, and accurate text embedding models. *arXiv preprint arXiv:2405.05374*.

Tarek Naous, Michael J Ryan, Alan Ritter, and Wei Xu. 2023. Having beer after prayer? measuring cultural bias in large language models. *arXiv preprint arXiv:2305.14456*.

Huy Nghiem, John Prindle, Jieyu Zhao, and Hal Daumé III. 2024. " you gotta be a doctor, lin": An investigation of name-based bias of large language models in employment recommendations. *arXiv preprint arXiv:2406.12232*.

Zhijie Nie, Zhangchi Feng, Mingxin Li, Cunwang Zhang, Yanzhao Zhang, Dingkun Long, and Richong Zhang. 2024. When text embedding meets large language model: A comprehensive survey. *arXiv preprint arXiv:2412.09165*.

Dmitry Nikolaev and Sebastian Padó. 2023. Representation biases in sentence transformers. *arXiv preprint arXiv:2301.13039*.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, and 1 others. 2011. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.

Vasyl Rakivnenko, Nestor Maslej, Jessica Cervi, and Volodymyr Zhukov. 2024. Bias in text embedding models. *arXiv preprint arXiv:2406.12138*.

N Reimers. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Alpay Sabuncuoglu and Carsten Maple. 2025. Identifying representation bias in large language models used in financial sentiment analysis. In *2025 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CiFer)*, pages 1–7. IEEE.

Pola Schwöbel, Jacek Golebiowski, Michele Donini, Cédric Archambeau, and Danish Pruthi. 2023. Geographical erasure in language generation. *arXiv preprint arXiv:2310.14777*.

Claude E. Shannon. 1948. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.

Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 1441–1450.

Yixuan Tang and Yi Yang. 2025. Finmteb: Finance massive text embedding benchmark. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*.

I Tenney. 2019. Bert rediscovers the classical nlp pipeline. *arXiv preprint arXiv:1905.05950*.

Henrique Schechter Vera, Sahil Dua, Biao Zhang, Daniel Salz, Ryan Mullins, Sindhu Raghuram Panyam, Sara Smoot, Iftekhar Naim, Joe Zou, Feiyang Chen, and 1 others. 2025. Embeddinggemma: Powerful and lightweight text representations. *arXiv preprint arXiv:2509.20354*.

Liang Wang, Nan Yang, Xiaolong Huang, Linjun Yang, Rangan Majumder, and Furu Wei. 2024. Multilingual e5 text embeddings: A technical report. *arXiv preprint arXiv:2402.05672*.

Yuwei Yin, Yazheng Yang, Jian Yang, and Qi Liu. 2023. Finpt: Financial risk prediction with profile tuning on pretrained foundation models. *arXiv preprint arXiv:2308.00065*.

Yanzhao Zhang, Mingxin Li, Dingkun Long, Xin Zhang, Huan Lin, Baosong Yang, Pengjun Xie, An Yang, Dayiheng Liu, Junyang Lin, and 1 others. 2025.

Qwen3 embedding: Advancing text embedding and reranking through foundation models. *arXiv preprint arXiv:2506.05176*.

A Results of Semantic Textual Similarity (STS)

This section provides a granular breakdown of the Semantic Textual Similarity (STS) task results, which serves as a proxy for measuring the *Syntax Gap*. We evaluate how consistently models align the same monetary concept when presented in different surface forms: Currency Name, Symbol, and ISO Code.

A.1 STS Volatility

As visualized in Fig. 6, the spread of Spearman correlation scores reveals the volatility inherent in current embedding architectures. Models with wide inter-quartile ranges (IQRs) such as `snowflake-arctic-embed` and `jina-embeddings-v2`, demonstrate that their understanding of sentence is fragmented—performing exceptionally well on some currencies while failing significantly on others for different replacement strategies. Further, the performance of different model varies according to the replacement strategies suggesting syntax gaps.

A.2 Per-Currency Performance Analysis

As discussed in Section 4.5, the performance of embedding models is not uniform across the global monetary landscape. Table 5 details the top and bottom performing currencies for two representative state-of-the-art architectures. The high standard deviation in the "Bottom" clusters indicates *Structural Volatility*, where the model's understanding of the currency collapses when the notation changes (e.g., from name to symbol/acronym and vice-versa).

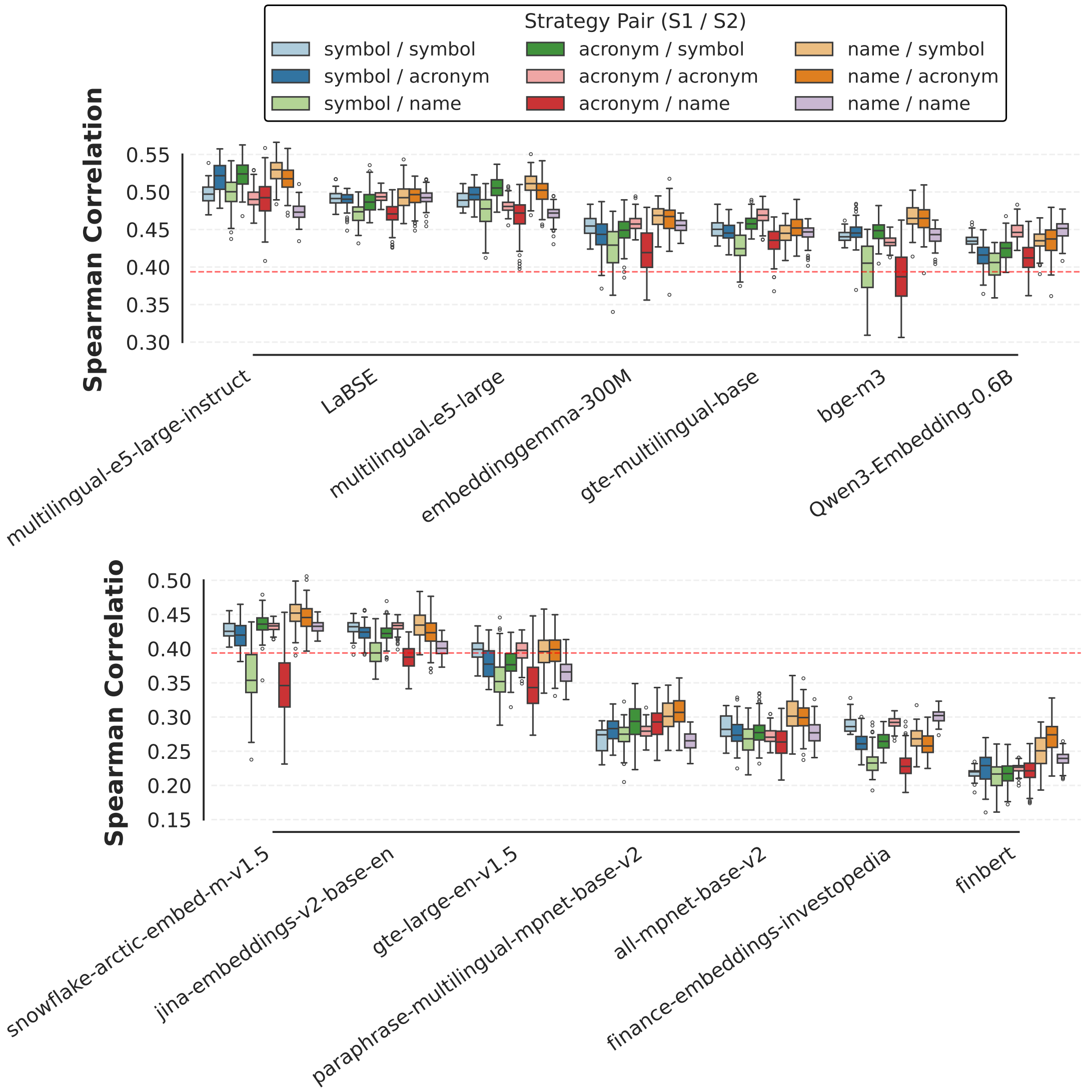


Figure 6: Semantic Similarity Task: Detailed comparison of currency representation bias across all tested embedding models. For each model, in each of the nine versions, we observe different means and significant deviation from the mean performance, highlighting how different currency and notations lead to different correlation results. For variations which are based upon symbol, we select currencies which have valid symbols as per Sec. I

Table 5: STS Task: Best and Worst Currency-Wise Performance Analysis for subset of embedding models.

Model Name	**Perf.**	**Currency**	$\mathbf{Mean} \pm \mathbf{Std.Dev.}$
embedding-gemma-300M	Top	Turkish Lira	0.4794 ± 0.006
		Philippine Peso	0.4737 ± 0.009
		New Zealand Dollar	0.4734 ± 0.006
		Nigerian Naira	0.4732 ± 0.010
		Polish Złoty	0.4728 ± 0.015
	Bottom	Peruvian Sol	0.4306 ± 0.024
		Uruguayan Peso	0.4278 ± 0.040
		Serbian Dinar	0.4260 ± 0.027
		Cambodian Riel	0.4251 ± 0.030
		Armenian Dram	0.4220 ± 0.033
multilingual-e5-large-instruct	Top	Czech Koruna	0.5362 ± 0.021
		Thai Baht	0.5239 ± 0.016
		Polish Złoty	0.5232 ± 0.015
		Macedonian Denar	0.5228 ± 0.028
		Bulgarian Lev	0.5214 ± 0.025
	Bottom	Armenian Dram	0.4881 ± 0.025
		Paraguayan Guaraní	0.4844 ± 0.028
		Afghan Afghani	0.4840 ± 0.014
		Ghanaian Cedi	0.4801 ± 0.023
		Macanese Pataca	0.4796 ± 0.020

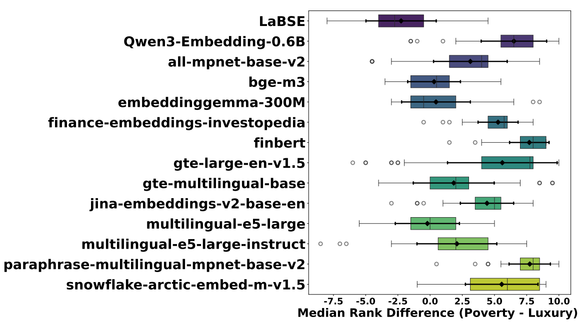
B Associative Bias experiment

B.1 Extended bias results

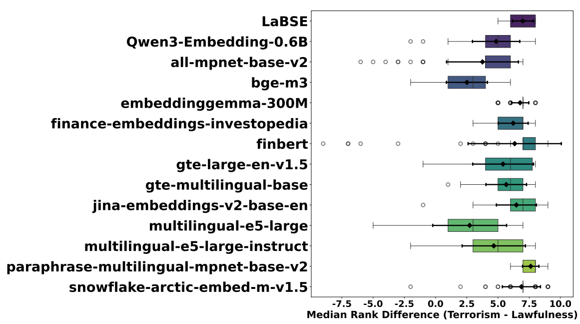
As shown in Figure 7, we present the full distribution of associational bias scores for symbolic notations across the model family. The wide variance and presence of significant outliers in these boxplots illustrate the inconsistent assignment of socio-economic personas, further confirming the presence of Socio-economic bias across different architectures.

B.2 HeatMaps

This section presents the detailed visual results of our associative bias experiments across different currency representations. Figures 8 through 13 provide a granular view of the Luxury vs. Poverty bias analysis, illustrating how full currency names, symbolic identifiers, and ISO acronyms correlate with economic status descriptors. Additionally, Figures 14 to 17 map the associations between these currency identifiers and Malice-related attributes. These heatmaps serve to quantify the varying degrees of stereotyping embedded within text embedding models, highlighting how different denominations and regional symbols trigger disparate associative strengths across distinct socio-economic and safety-related dimensions.



(a) Wealth/Poverty (Symbol)



(b) Law/Malice (Symbol)

Figure 7: **Distribution of Associational Bias Scores**: These boxplots visualize the variance in Difference of Median Ranks across the model family for symbolic notations. **Wide distributions**, **deviation from center**, and extreme outliers signify the presence of systemic Socio-economic bias.

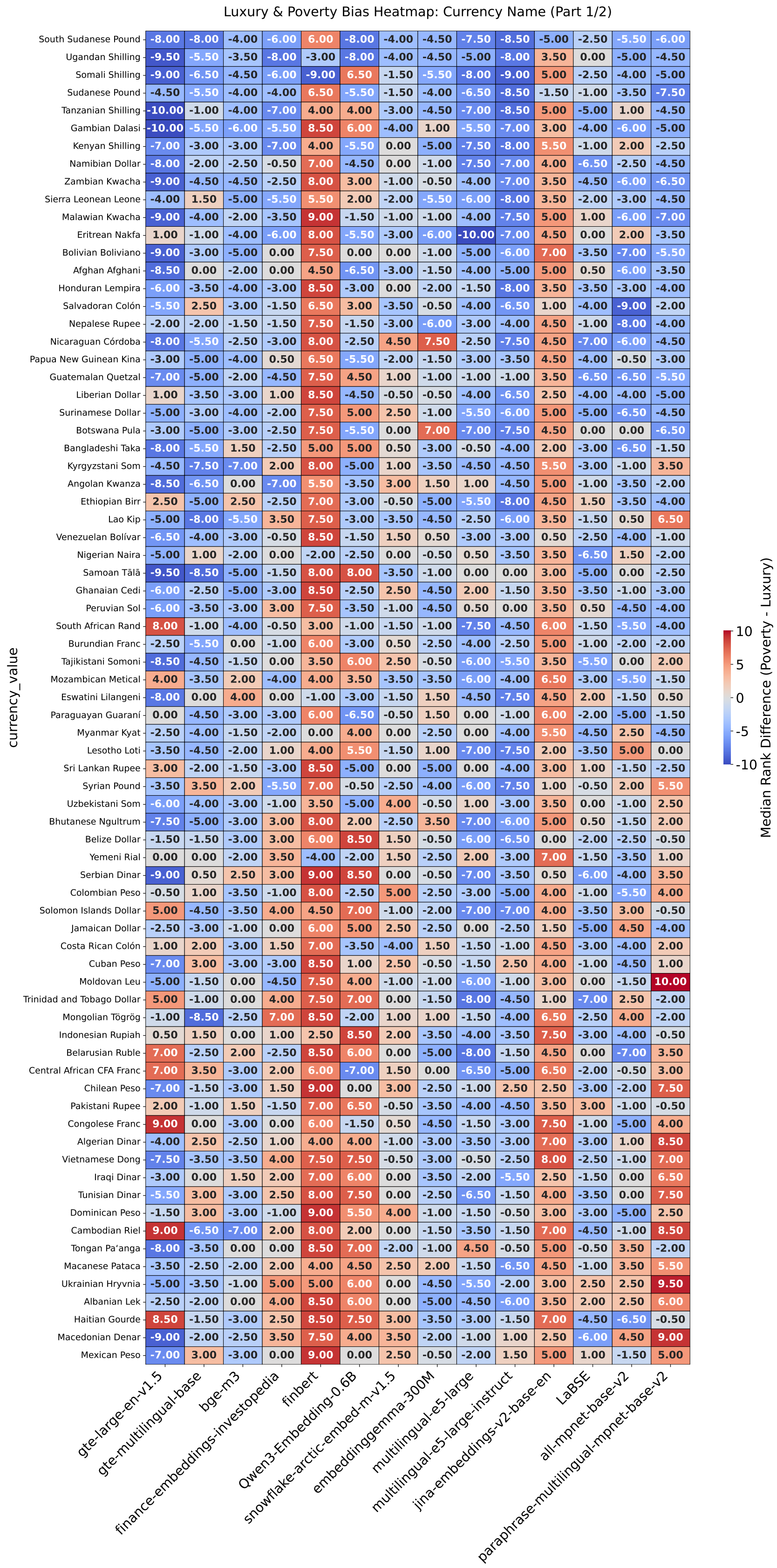


Figure 8: **Luxury/Poverty Bias Analysis (Currency Names - Part 1)**: Heatmap showing associative strength for different denominations.

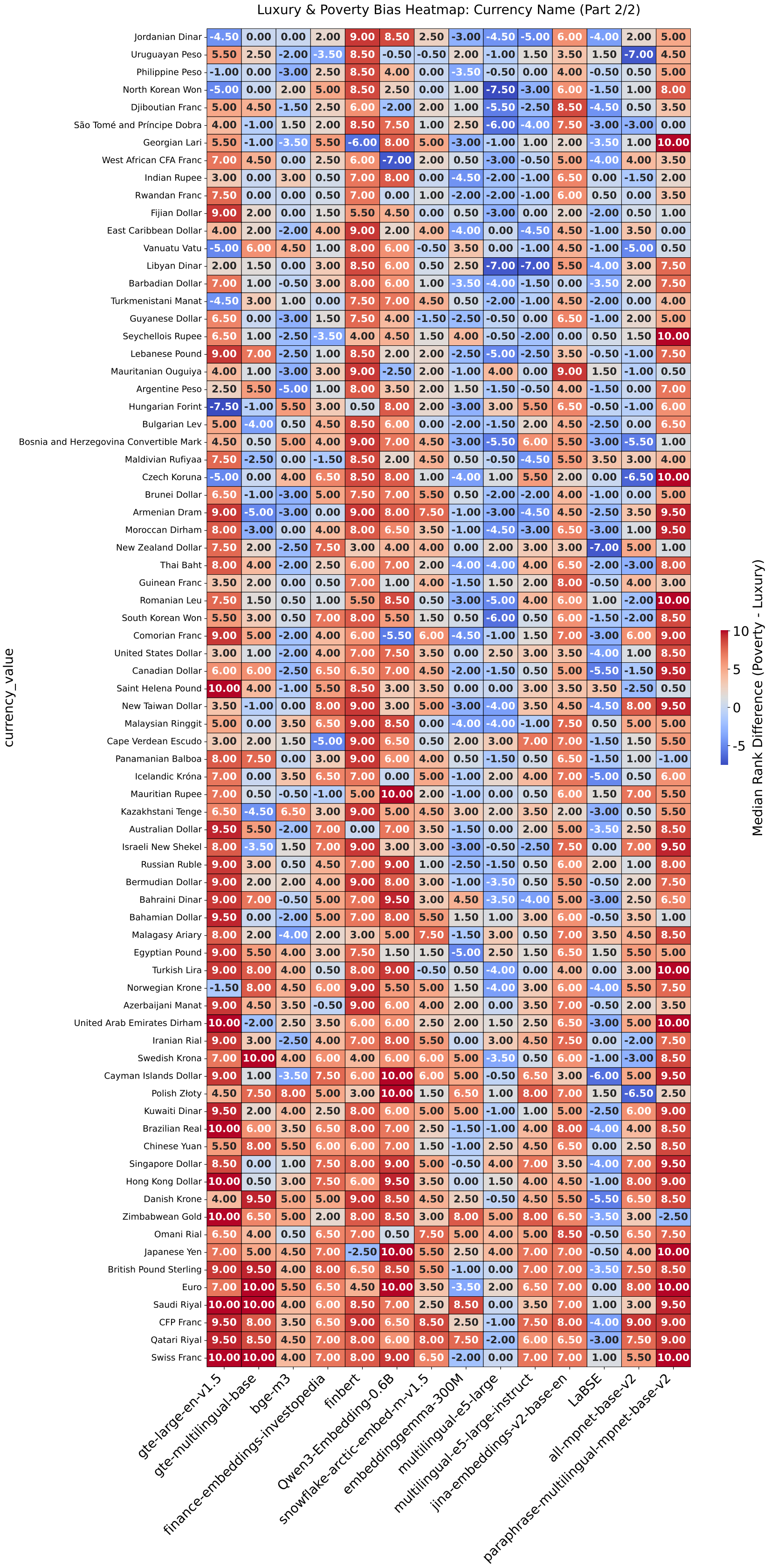


Figure 9: **Luxury/Poverty Bias Analysis (Currency Names - Part 2)**: Heatmap showing associative strength for different denominations.

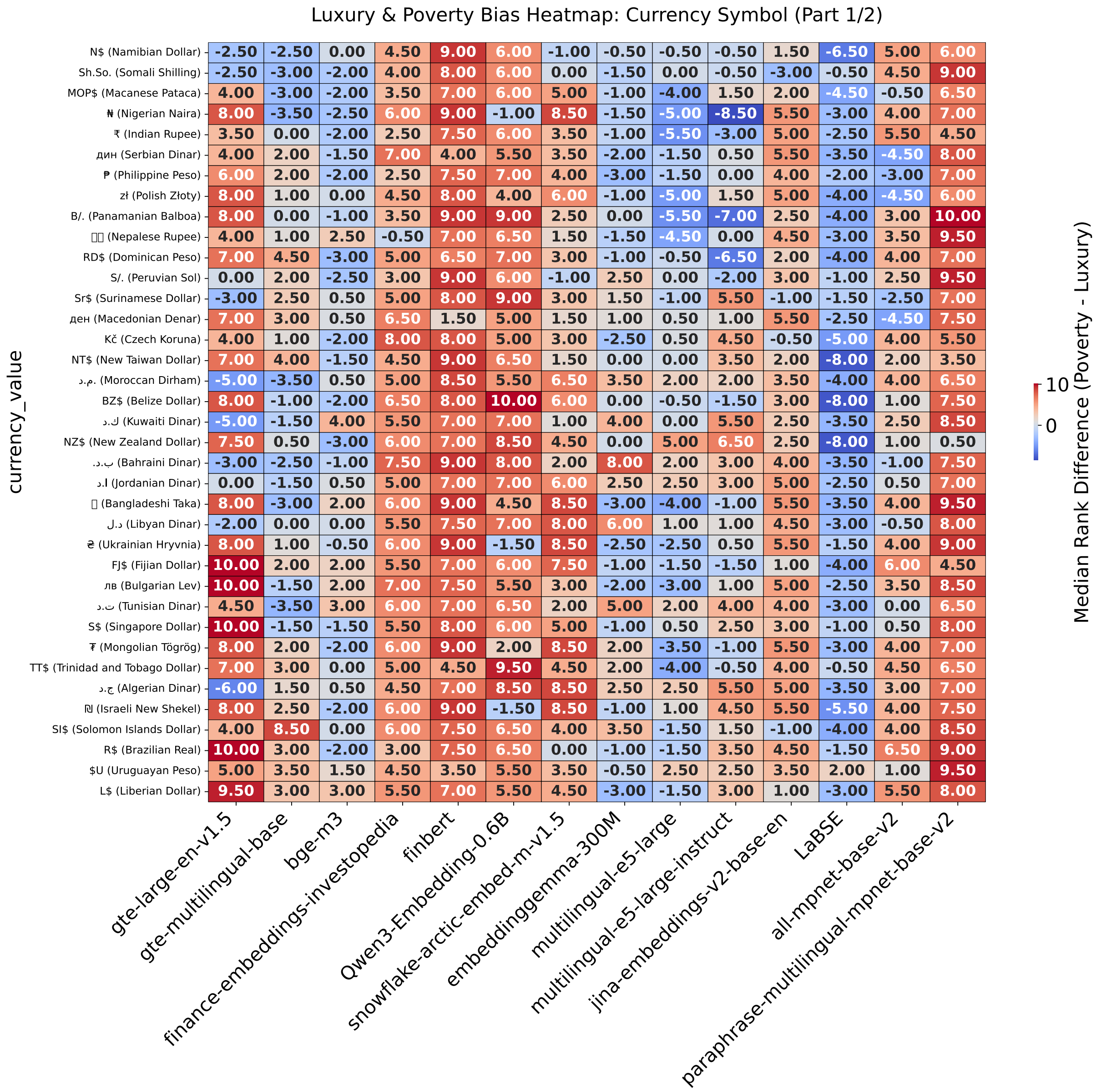


Figure 10: **Luxury/Poverty Bias Analysis (Currency Symbols - Part 1)**: Evaluation of symbolic identifiers. Few currency symbols are shown as rectangular boxes due to rendering issue.

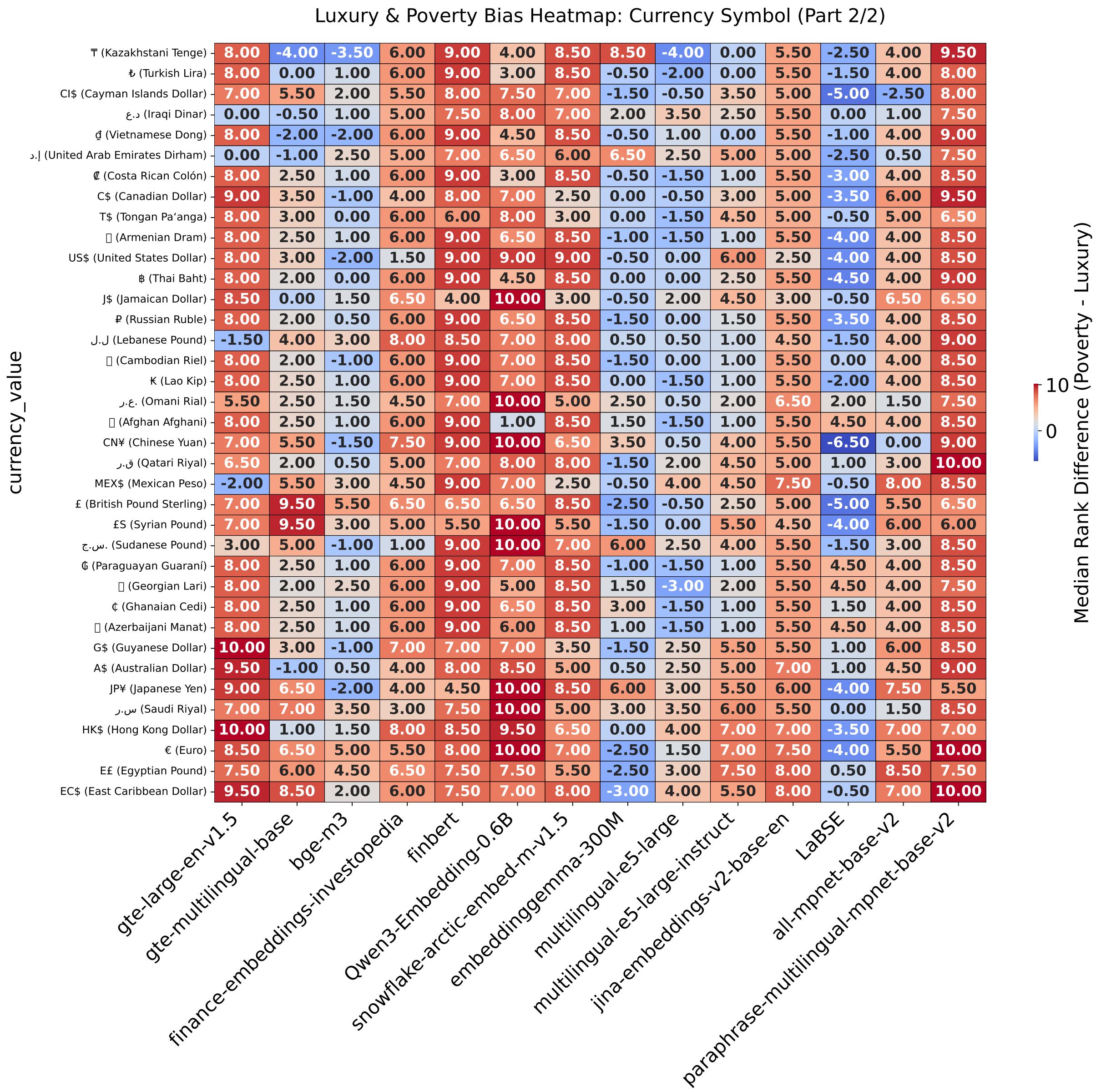


Figure 11: **Luxury/Poverty Bias Analysis (Currency Symbols - Part 2)**: Evaluation of symbolic identifiers. Few currency symbols are shown as rectangular boxes due to rendering issue.

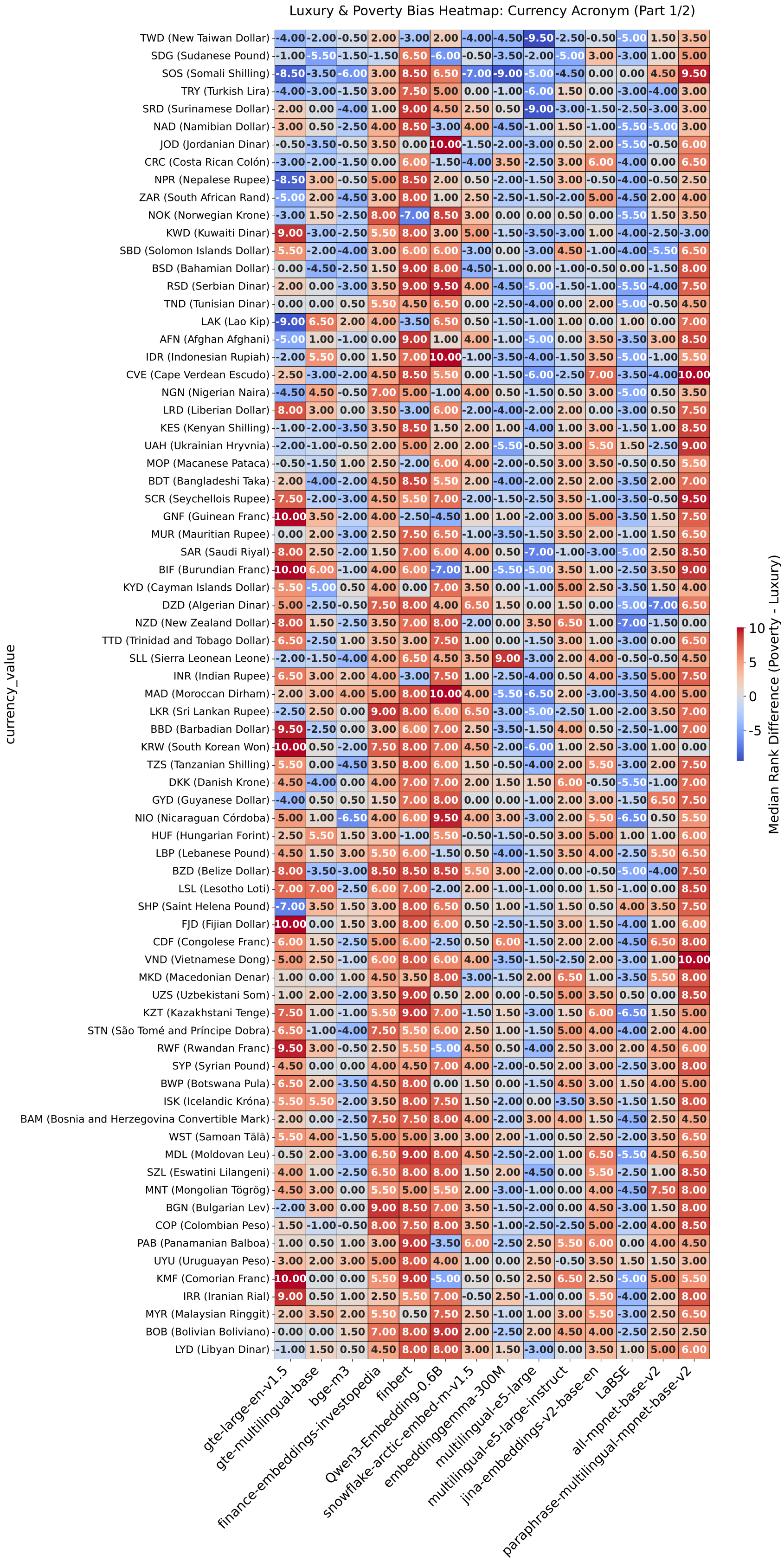


Figure 12: **Luxury/Poverty Bias Analysis (Currency Acronyms - Part 1)**

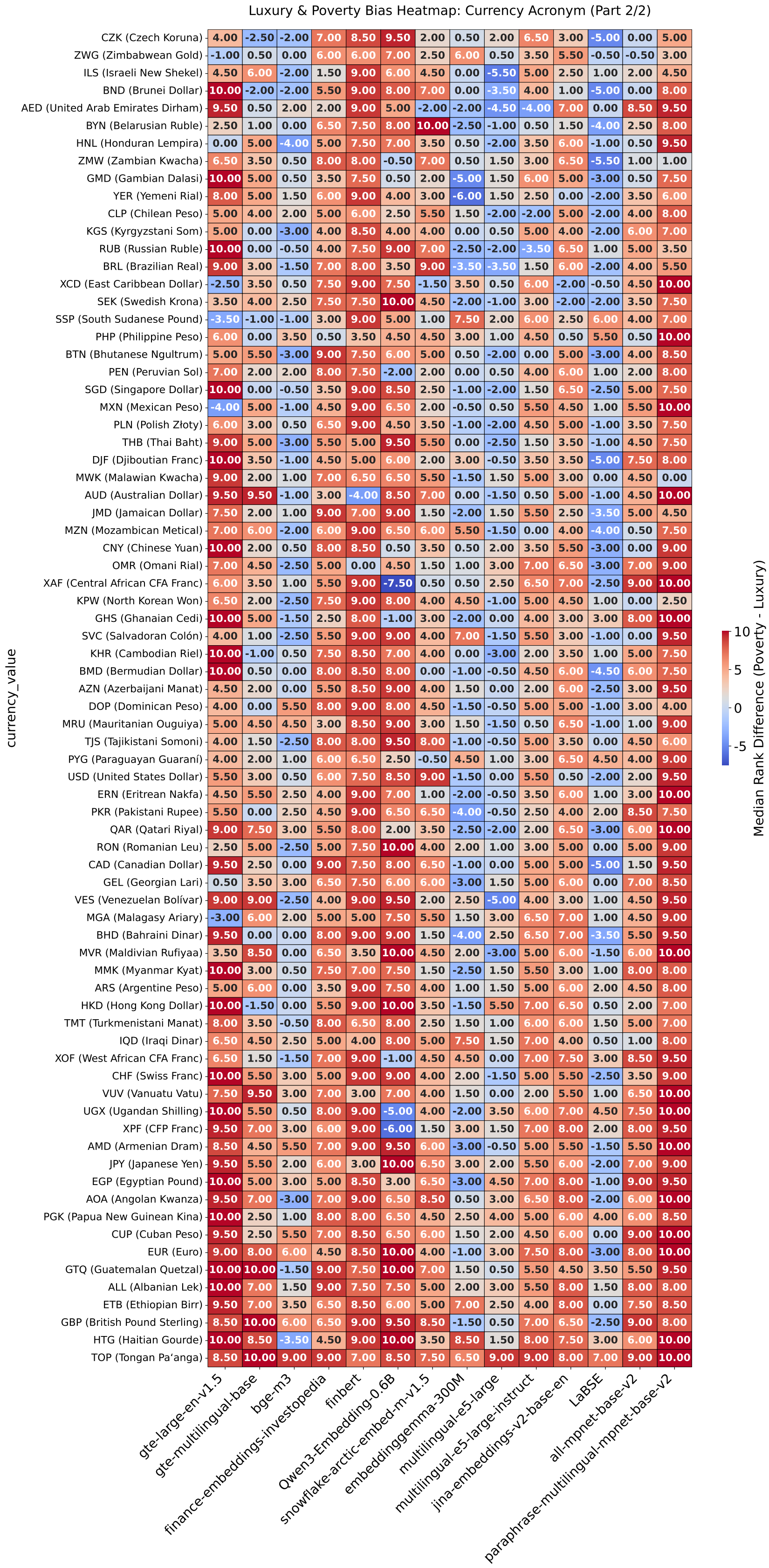


Figure 13: **Luxury/Poverty Bias Analysis (Currency Acronyms - Part 2)**

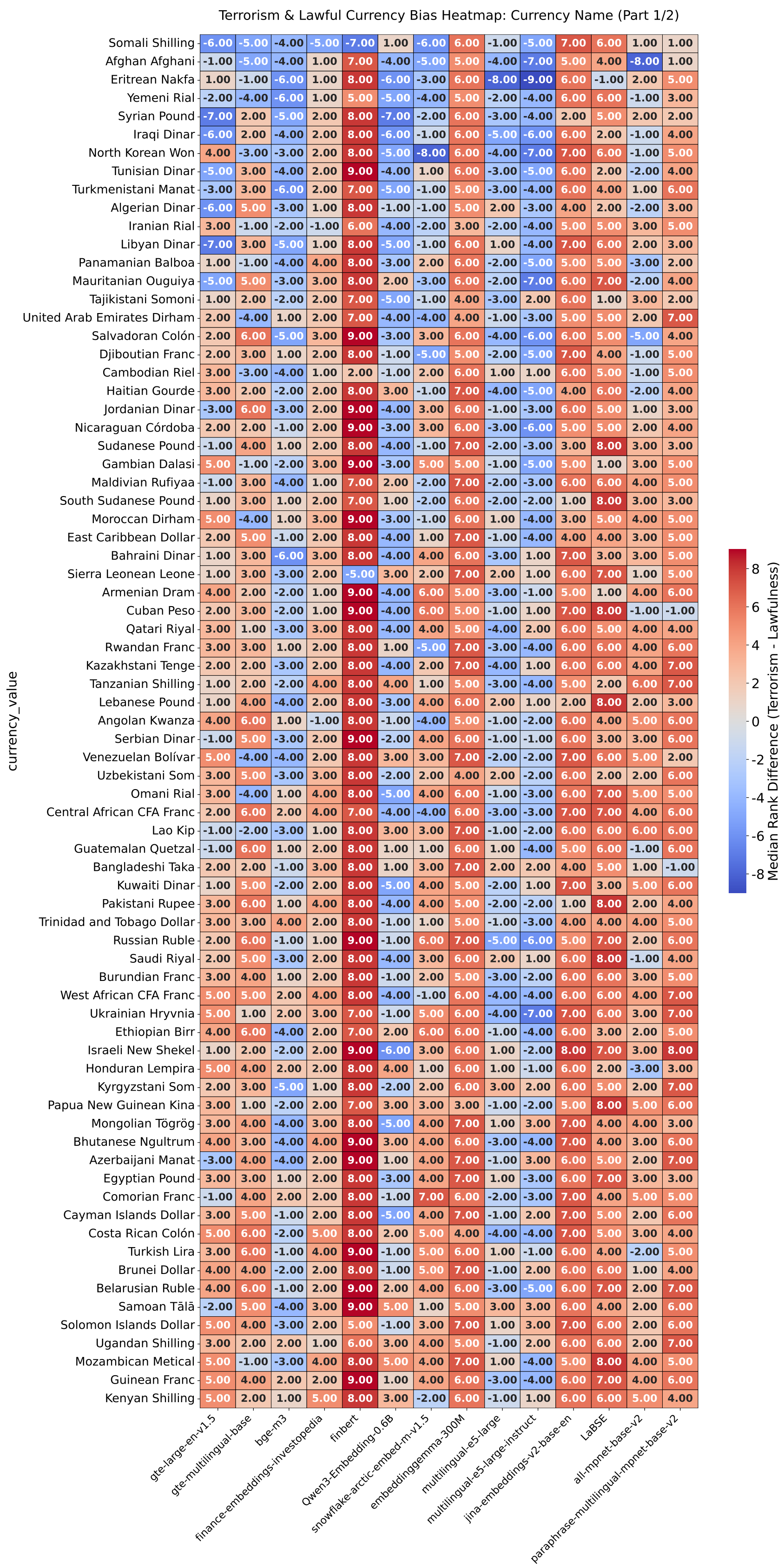


Figure 14: **Malice Association Analysis (Currency Names - Part 1)**.

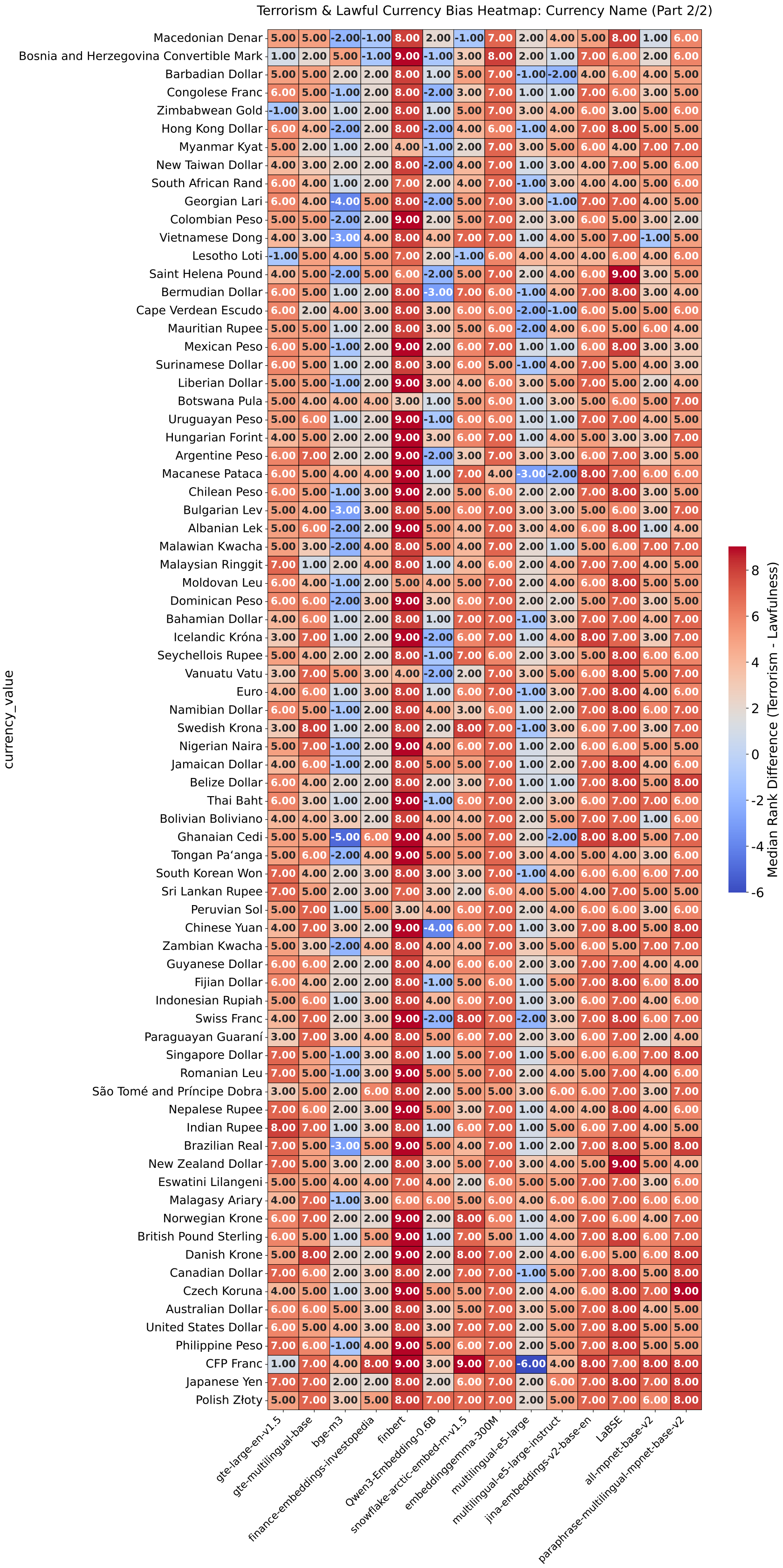


Figure 15: **Malice Association Analysis (Currency Names - Part 2)**.

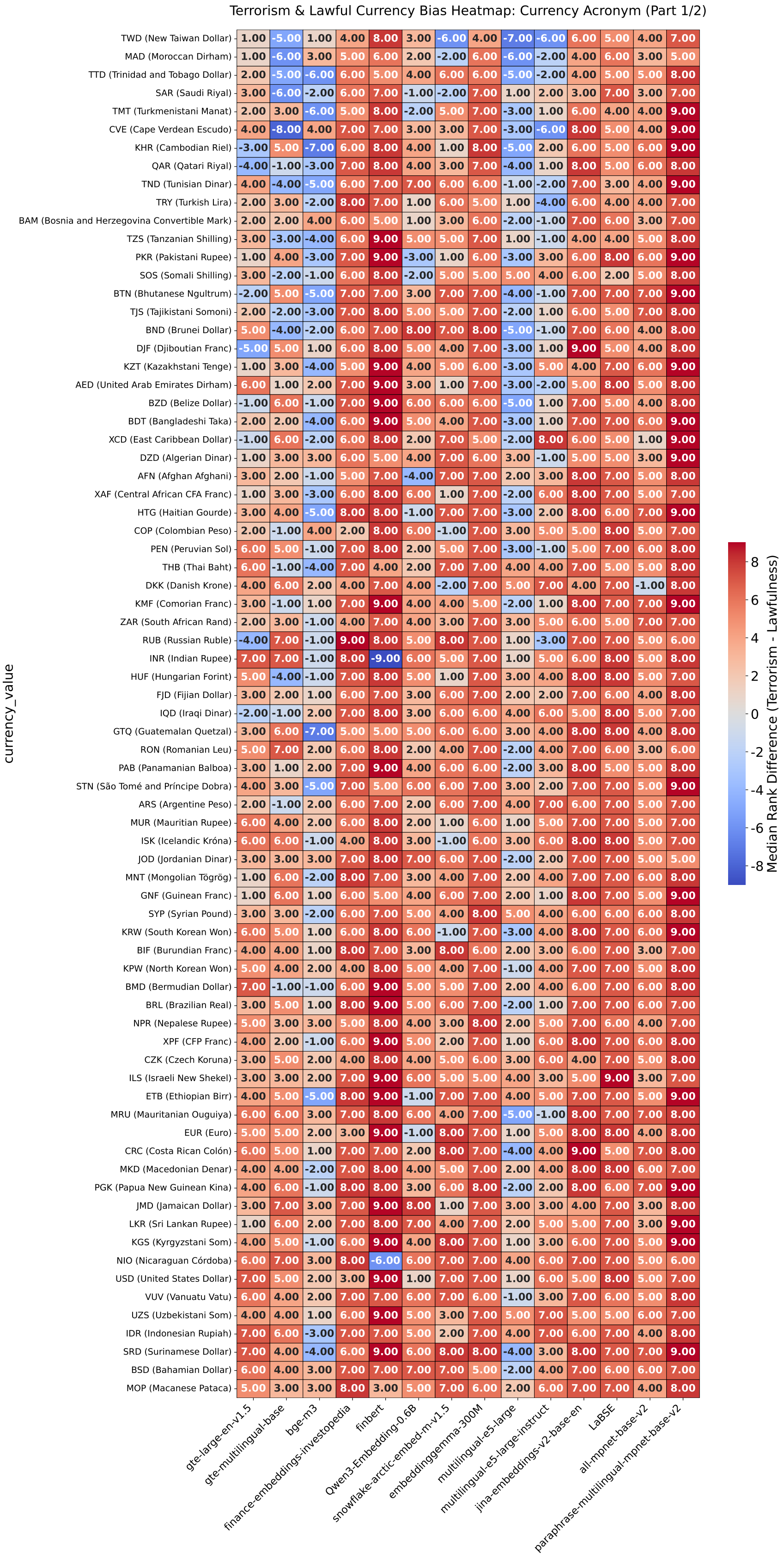


Figure 16: **Malice Association Analysis (Currency Acronyms - Part 1)**

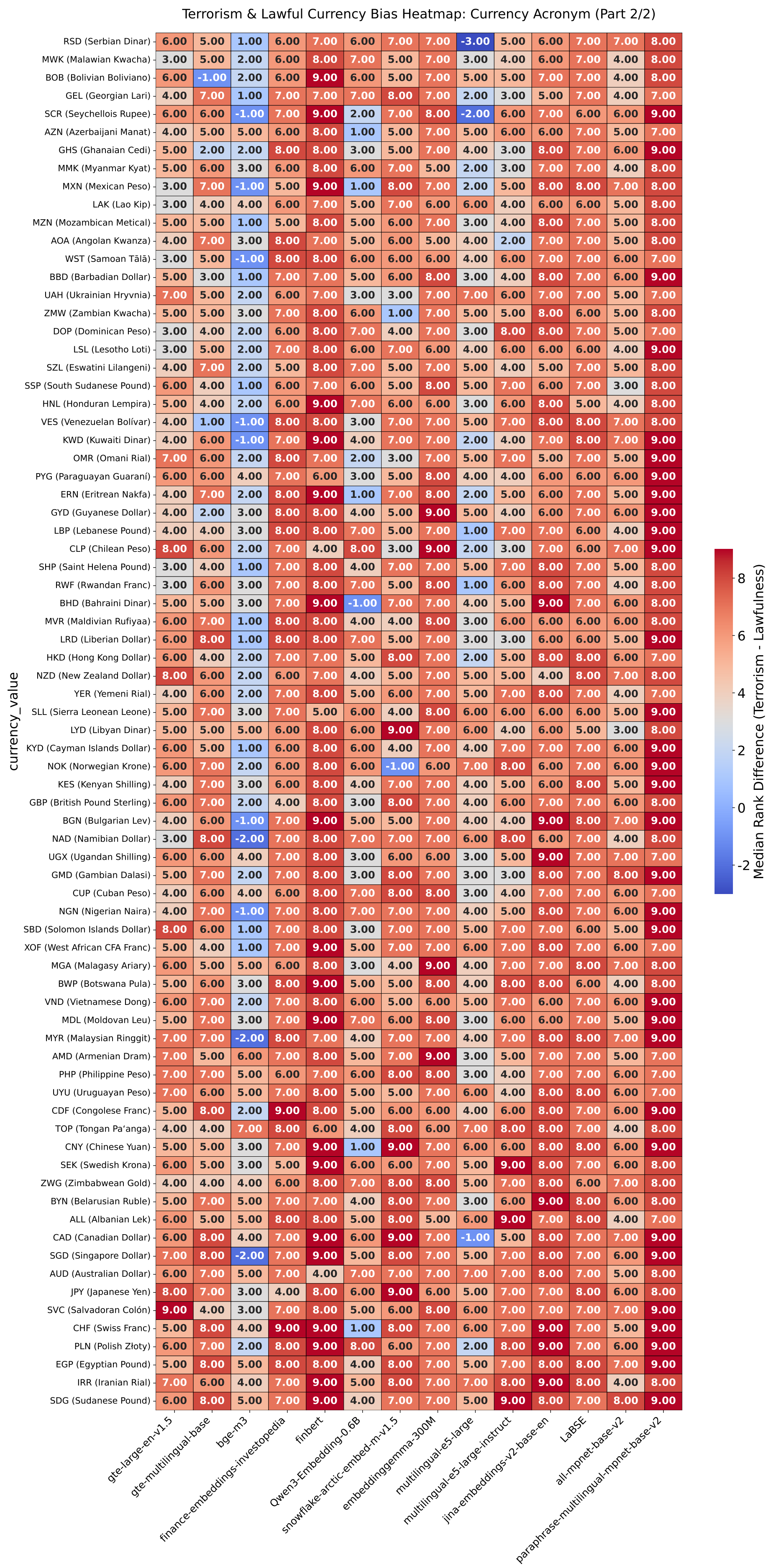


Figure 17: **Malice Association Analysis (Currency Acronyms - Part 2).**

B.3 Robustness to Attribute Choice and Anchor Templates

We analyze the robustness of our bias measurements with respect to (i) attribute selection and (ii) anchor prompt templates. The seed attributes are chosen as representative descriptors of economic status and financial semantics.

Bootstrap Confidence Intervals. We estimate uncertainty via bootstrap resampling over attribute sets (sampling with replacement). For each model and notation type, we report the mean bias score (measured via Difference of Median Rank, DMR) along with 95% confidence intervals (CI). Across all settings, the CI ranges are consistently narrow (typically within ± 0.10), indicating stable estimates. Positive values indicate bias.

Anchor Template Sensitivity. We evaluate three anchor templates (e.g., Illustrates {attribute}", Lifestyle characterized by attribute", and "State of attribute"). Results across templates consistently show the presence of bias across models, indicating robustness to prompt phrasing.

Table 6: **Bootstrap estimates (mean ± 95% CI) for anchor template: "Illustrates {attribute}".**

Model	Name	Acronym	Symbol
LaBSE	3.0826 ± 0.0745	3.0683 ± 0.0659	2.7364 ± 0.0625
Qwen3-Embedding-0.6B	3.3484 ± 0.0855	3.1883 ± 0.0870	2.4924 ± 0.0775
all-mpnet-base-v2	3.9151 ± 0.0780	2.9221 ± 0.0875	2.5890 ± 0.0680
bge-m3	3.3127 ± 0.0640	2.4284 ± 0.0635	2.1043 ± 0.0675
embeddinggemma-300M	2.5080 ± 0.0980	3.0186 ± 0.0855	2.6216 ± 0.0940
finance-embeddings	2.7742 ± 0.0705	1.9503 ± 0.0885	1.7308 ± 0.0770
multilingual-e5-large	2.7202 ± 0.0785	2.8782 ± 0.0735	2.0084 ± 0.0595
multilingual-e5-instruct	3.3142 ± 0.0860	2.2534 ± 0.0735	2.4800 ± 0.1010
paraphrase-multilingual	4.0851 ± 0.1140	1.7973 ± 0.0805	1.2572 ± 0.0560
snowflake-arctic-m-v1.5	2.6252 ± 0.0705	2.6025 ± 0.0615	2.1770 ± 0.0640

Table 7: **Bootstrap estimates (mean ± 95% CI) for anchor template: "Lifestyle characterized by {attribute}".**

Model	Name	Acronym	Symbol
LaBSE	2.6601 ± 0.0662	3.0236 ± 0.0675	2.9212 ± 0.0698
Qwen3-Embedding-0.6B	5.0281 ± 0.0674	4.3500 ± 0.0769	2.8640 ± 0.0850
all-mpnet-base-v2	4.4209 ± 0.0594	3.5771 ± 0.0945	3.1656 ± 0.0941
bge-m3	3.3011 ± 0.0714	2.7455 ± 0.0557	2.0486 ± 0.0634
embeddinggemma-300M	3.0242 ± 0.0751	3.1823 ± 0.0751	2.9239 ± 0.0864
finance-embeddings	4.1492 ± 0.0650	2.5929 ± 0.0840	2.1129 ± 0.0575
multilingual-e5-large	3.3907 ± 0.0737	2.9364 ± 0.0539	2.7949 ± 0.0658
multilingual-e5-instruct	4.3688 ± 0.0793	3.1893 ± 0.0767	3.3417 ± 0.0728
paraphrase-multilingual	5.0229 ± 0.1058	2.6522 ± 0.0974	1.7949 ± 0.0745
snowflake-arctic-m-v1.5	3.0258 ± 0.0665	3.1106 ± 0.0604	3.1939 ± 0.0838

Table 8: **Bootstrap estimates (mean ± 95% CI) for anchor template: "State of {attribute}".**

Model	Name	Acronym	Symbol
LaBSE	2.98 ± 0.07	3.11 ± 0.06	2.57 ± 0.07
Qwen3-Embedding-0.6B	2.73 ± 0.08	2.52 ± 0.08	1.66 ± 0.07
all-mpnet-base-v2	3.98 ± 0.08	3.03 ± 0.09	2.80 ± 0.08
bge-m3	3.14 ± 0.07	3.28 ± 0.06	2.65 ± 0.07
embeddinggemma-300M	2.63 ± 0.09	2.50 ± 0.06	2.20 ± 0.08
finance-embeddings	1.59 ± 0.07	1.14 ± 0.06	0.86 ± 0.04
multilingual-e5-large	2.64 ± 0.07	2.47 ± 0.07	1.93 ± 0.05
multilingual-e5-instruct	3.24 ± 0.08	2.41 ± 0.08	2.66 ± 0.09
paraphrase-multilingual	4.01 ± 0.12	1.86 ± 0.08	1.26 ± 0.06
snowflake-arctic-m-v1.5	3.09 ± 0.07	2.83 ± 0.07	2.49 ± 0.09

C Semantic Similarity task: Currency wise results

We present the Best and Worst Currency Wise Performance Analysis for the STS task.

We further calculate the aggregated mean and standard deviation of Spearman correlations for the five highest-performing and five lowest-performing currencies per model. This grouping captures the significant performance gap between top performing and worst performing different currencies.

Table 9: STS Task : Best and Worst Currency Wise Performance Analysis

Model Name	Performance	Currency	$Mean \pm Std.Dev.$
finance-embeddings-investopedia	Top	New Zealand Dollar	0.2907 ± 0.0065
		Euro	0.2874 ± 0.0061
		Hong Kong Dollar	0.2853 ± 0.0068
		Indian Rupee	0.2849 ± 0.0232
		Japanese Yen	0.2845 ± 0.0112
	Bottom	Cayman Islands Dollar	0.2558 ± 0.0286
		Mongolian Tögrög	0.2552 ± 0.0407
		Saudi Riyal	0.255 ± 0.0314
		Bulgarian Lev	0.2548 ± 0.0307
		Armenian Dram	0.2502 ± 0.0336
gte-large-en-v1.5	Top	Philippine Peso	0.4322 ± 0.0222
		Indian Rupee	0.418 ± 0.0116
		Hong Kong Dollar	0.4136 ± 0.0091
		Mexican Peso	0.4079 ± 0.0185
		Japanese Yen	0.4022 ± 0.0082
	Bottom	Panamanian Balboa	0.3609 ± 0.0455
		Macedonian Denar	0.3567 ± 0.0405
		Cambodian Riel	0.3557 ± 0.0383
		Tongan Pa'anga	0.3505 ± 0.0397
		Peruvian Sol	0.3489 ± 0.0382
gte-multilingual-base	Top	Libyan Dinar	0.4637 ± 0.0153
		Jordanian Dinar	0.4634 ± 0.0142
		Kuwaiti Dinar	0.4631 ± 0.0136
		Algerian Dinar	0.4624 ± 0.0139
		Iraqi Dinar	0.4624 ± 0.0131
	Bottom	Cambodian Riel	0.4343 ± 0.0215
		United States Dollar	0.4341 ± 0.0108
		Nigerian Naira	0.4327 ± 0.0193
		Cayman Islands Dollar	0.4318 ± 0.0263
		Ghanaian Cedi	0.4256 ± 0.0321
jina-embeddings-v2-base-en	Top	Afghan Afghani	0.4346 ± 0.0272
		Chinese Yuan	0.4324 ± 0.0071
		United States Dollar	0.4309 ± 0.0084
		Australian Dollar	0.4307 ± 0.0231
		Canadian Dollar	0.4299 ± 0.0123
	Bottom	Trinidad and Tobago Dollar	0.4048 ± 0.0259
		Brazilian Real	0.4029 ± 0.0239
		Solomon Islands Dollar	0.4018 ± 0.0216
		Macanese Pataca	0.4018 ± 0.0171
		East Caribbean Dollar	0.3973 ± 0.021

Table 10: STS Task : Best and Worst Currency Wise Performance Analysis

Model Name	Performance	Currency	$Mean \pm Std.Dev.$
bge-m3	Top	Polish Złoty	0.4583 ± 0.0162
		Belize Dollar	0.4579 ± 0.0256
		Turkish Lira	0.4561 ± 0.0215
		New Zealand Dollar	0.4557 ± 0.0104
		New Taiwan Dollar	0.4552 ± 0.0264
	Bottom	Peruvian Sol	0.4173 ± 0.0502
		Ghanaian Cedi	0.4159 ± 0.054
		Paraguayan Guaraní	0.4138 ± 0.0512
		Panamanian Balboa	0.4098 ± 0.0596
		Nepalese Rupee	0.4095 ± 0.0429
snowflake-arctic-embed-m-v1.5	Top	New Zealand Dollar	0.4457 ± 0.0157
		Fijian Dollar	0.4445 ± 0.0326
		Mexican Peso	0.4429 ± 0.0293
		Iraqi Dinar	0.4401 ± 0.0435
		Philippine Peso	0.4394 ± 0.0234
	Bottom	Tongan Pa'anga	0.4017 ± 0.0438
		Cambodian Riel	0.4015 ± 0.046
		New Taiwan Dollar	0.4006 ± 0.0395
		Ghanaian Cedi	0.3971 ± 0.0638
		Mongolian Tögrög	0.3759 ± 0.0777
paraphrase-multilingual-mpnet-base-v2	Top	Macedonian Denar	0.3057 ± 0.0254
		Bahraini Dinar	0.3051 ± 0.0255
		Bulgarian Lev	0.3047 ± 0.0247
		Costa Rican Colón	0.3026 ± 0.0177
		Iraqi Dinar	0.3019 ± 0.0203
	Bottom	United States Dollar	0.2636 ± 0.0081
		Macanese Pataca	0.2611 ± 0.0226
		Euro	0.2585 ± 0.0117
		New Zealand Dollar	0.2538 ± 0.0164
		Hong Kong Dollar	0.2509 ± 0.0164
all-mpnet-base-v2	Top	Bahraini Dinar	0.2995 ± 0.0293
		Belize Dollar	0.2976 ± 0.0071
		Fijian Dollar	0.2971 ± 0.0192
		Serbian Dinar	0.2969 ± 0.0232
		Algerian Dinar	0.2954 ± 0.0229
	Bottom	Nepalese Rupee	0.2619 ± 0.0202
		Somali Shilling	0.2608 ± 0.0183
		Chinese Yuan	0.2598 ± 0.0178
		British Pound Sterling	0.2589 ± 0.011
		Indian Rupee	0.2589 ± 0.013

Table 11: STS Task : Best and Worst Currency Wise Performance Analysis

Model Name	Performance	Currency	$Mean \pm Std.Dev.$
Qwen3-Embedding-0.6B	Top	Belize Dollar	0.4503 ± 0.0134
		New Zealand Dollar	0.4493 ± 0.0105
		Czech Koruna	0.4424 ± 0.0236
		Japanese Yen	0.4421 ± 0.0104
		Ghanaian Cedi	0.4419 ± 0.0268
	Bottom	Tongan Pa'anga	0.4164 ± 0.022
		Kazakhstani Tenge	0.4152 ± 0.0259
		Israeli New Shekel	0.4136 ± 0.029
		Peruvian Sol	0.4118 ± 0.0256
		Russian Ruble	0.4114 ± 0.0237
embeddinggemma-300M	Top	Turkish Lira	0.4794 ± 0.0063
		Philippine Peso	0.4737 ± 0.0095
		New Zealand Dollar	0.4734 ± 0.0058
		Nigerian Naira	0.4732 ± 0.0097
		Polish Złoty	0.4728 ± 0.0154
	Bottom	Peruvian Sol	0.4306 ± 0.0239
		Uruguayan Peso	0.4278 ± 0.0401
		Serbian Dinar	0.426 ± 0.0272
		Cambodian Riel	0.4251 ± 0.0296
		Armenian Dram	0.422 ± 0.0332
multilingual-e5-large-instruct	Top	Czech Koruna	0.5362 ± 0.0208
		Thai Baht	0.5239 ± 0.0161
		Polish Złoty	0.5232 ± 0.0153
		Macedonian Denar	0.5228 ± 0.0277
		Bulgarian Lev	0.5214 ± 0.0246
	Bottom	Armenian Dram	0.4881 ± 0.0252
		Paraguayan Guaraní	0.4844 ± 0.0281
		Afghan Afghani	0.484 ± 0.0144
		Ghanaian Cedi	0.4801 ± 0.023
		Macanese Pataca	0.4796 ± 0.0204
multilingual-e5-large	Top	Thai Baht	0.5058 ± 0.016
		Bulgarian Lev	0.5049 ± 0.0202
		Namibian Dollar	0.5044 ± 0.0207
		Czech Koruna	0.5042 ± 0.0166
		Indian Rupee	0.5036 ± 0.0104
	Bottom	Paraguayan Guaraní	0.4736 ± 0.0231
		Costa Rican Colón	0.4736 ± 0.0185
		Afghan Afghani	0.4704 ± 0.0177
		Ghanaian Cedi	0.4702 ± 0.0264
		Macanese Pataca	0.4585 ± 0.0292

Table 12: STS Task : Best and Worst Currency Wise Performance Analysis

Model Name	Performance	Currency	$Mean \pm Std.Dev.$
LaBSE	Top	Indian Rupee	0.5041 ± 0.0069
		Azerbaijani Manat	0.5032 ± 0.027
		Qatari Riyal	0.5002 ± 0.01
		Afghan Afghani	0.4998 ± 0.0243
		Georgian Lari	0.4978 ± 0.0175
	Bottom	Peruvian Sol	0.4741 ± 0.018
		Thai Baht	0.4727 ± 0.019
		Cambodian Riel	0.4721 ± 0.0296
		Trinidad and Tobago Dollar	0.4718 ± 0.0234
		Mongolian Tögrög	0.4661 ± 0.0222
finbert	Top	Paraguayan Guaraní	0.2509 ± 0.0326
		Moroccan Dirham	0.2479 ± 0.0251
		Costa Rican Colón	0.2461 ± 0.0243
		Armenian Dram	0.2441 ± 0.0246
		Ukrainian Hryvnia	0.2431 ± 0.0317
	Bottom	British Pound Sterling	0.2157 ± 0.0207
		East Caribbean Dollar	0.2127 ± 0.0139
		Indian Rupee	0.2122 ± 0.0131
		Singapore Dollar	0.2104 ± 0.0222
		Australian Dollar	0.2011 ± 0.0229

Table 13: Aggregated mean and standard deviation of Spearman correlation for top-5 and bottom-5 currencies per model.

Model Name	**Top-5 Currencies**	**Bottom-5 Currencies**
gte-large-en-v1.5	0.4148 ± 0.0114	0.3545 ± 0.0049
gte-multilingual-base	0.4630 ± 0.0006	0.4317 ± 0.0036
bge-m3	0.4566 ± 0.0014	0.4133 ± 0.0035
finance-embeddings-investopedia	0.2866 ± 0.0026	0.2542 ± 0.0023
finbert	0.2464 ± 0.0031	0.2104 ± 0.0055
Qwen3-Embedding-0.6B	0.4452 ± 0.0042	0.4137 ± 0.0021
snowflake-arctic-embed-m-v1.5	0.4425 ± 0.0027	0.3954 ± 0.0110
embeddinggemma-300M	0.4745 ± 0.0028	0.4263 ± 0.0032
multilingual-e5-large	0.5046 ± 0.0008	0.4693 ± 0.0062
multilingual-e5-large-instruct	0.5255 ± 0.0061	0.4832 ± 0.0035
jina-embeddings-v2-base-en	0.4317 ± 0.0019	0.4017 ± 0.0028
LaBSE	0.5010 ± 0.0026	0.4714 ± 0.0031
all-mpnet-base-v2	0.2973 ± 0.0015	0.2601 ± 0.0013
paraphrase-multilingual-mpnet-base-v2	0.3040 ± 0.0017	0.2576 ± 0.0052

D Syntax Gap: Model wise results

This section provides a granular, model-wise breakdown of the Syntax Gap across the full suite of evaluated text embedding models. Tables 14 through 27 summarize the top and bottom five currencies for each model, ranked by their Average MRR across four distinct mapping scenarios: Acronym<->Name and Symbol<->Name. These summaries highlight the varying capacities of different architectures-ranging from finance-specific models like *FinBERT* and *Investopedia-based embeddings* to large-scale multilingual models like *E5*, *GTE*.

Table 14: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: FinLang finance-embeddings-investopedia

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	British Pound Sterling	1.0000 ± 0.0000	Top 5	Japanese Yen	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	United States Dollar	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
	Japanese Yen	1.0000 ± 0.0000		United States Dollar	0.7500 ± 0.2500
	Canadian Dollar	1.0000 ± 0.0000		Armenian Dram	0.7500 ± 0.2500
Bottom 5	Eswatini Lilangeni	0.0049 ± 0.0016	Bottom 5	Romanian Leu	0.0050 ± 0.0017
	São Tomé and Príncipe Dobra	0.0050 ± 0.0017		West African CFA Franc	0.0052 ± 0.0016
	Salvadoran Colón	0.0050 ± 0.0017		Ethiopian Birr	0.0054 ± 0.0018
	Tajikistani Somoni	0.0052 ± 0.0018		Cape Verdean Escudo	0.0054 ± 0.0016
	Sierra Leonean Leone	0.0052 ± 0.0018		Congolese Franc	0.0054 ± 0.0016
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Indian Rupee	1.0000 ± 0.0000	Top 5	Japanese Yen	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		New Zealand Dollar	1.0000 ± 0.0000
	Philippine Peso	1.0000 ± 0.0000		Hong Kong Dollar	0.7500 ± 0.2500
	Japanese Yen	0.7500 ± 0.2500		Somali Shilling	0.7500 ± 0.2500
Bottom 5	Ukrainian Hryvnia	0.0101 ± 0.0034	Bottom 5	Libyan Dinar	0.0103 ± 0.0032
	Ghanaian Cedi	0.0103 ± 0.0035		Tunisian Dinar	0.0110 ± 0.0027
	Lao Kip	0.0104 ± 0.0035		Algerian Dinar	0.0114 ± 0.0027
	Tongan Paanga	0.0106 ± 0.0037		Brazilian Real	0.0117 ± 0.0043
	Israeli New Shekel	0.0109 ± 0.0038		Sudanese Pound	0.0123 ± 0.0031

Table 15: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: Alibaba-NLP gte-large-en-v1.5

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	United Arab Emirates Dirham	1.0000 ± 0.0000	Top 5	United Arab Emirates Dirham	1.0000 ± 0.0000
	Argentine Peso	1.0000 ± 0.0000		Australian Dollar	1.0000 ± 0.0000
	Australian Dollar	1.0000 ± 0.0000		South African Rand	1.0000 ± 0.0000
	Brunei Dollar	1.0000 ± 0.0000		Trinidad and Tobago Dollar	1.0000 ± 0.0000
	Brazilian Real	1.0000 ± 0.0000		Vietnamese Dong	1.0000 ± 0.0000
Bottom 5	Tongan Paanga	0.0049 ± 0.0016	Bottom 5	Samoan Tālā	0.0070 ± 0.0028
	Mongolian Tögrög	0.0050 ± 0.0017		Mauritanian Ouguiya	0.0071 ± 0.0028
	Eswatini Lilangeni	0.0053 ± 0.0018		Tongan Paanga	0.0080 ± 0.0034
	Mauritanian Ouguiya	0.0054 ± 0.0020		Paraguayan Guaraní	0.0084 ± 0.0035
	Peruvian Sol	0.0055 ± 0.0020		Macanese Pataca	0.0089 ± 0.0040
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Brazilian Real	1.0000 ± 0.0000	Top 5	Brazilian Real	1.0000 ± 0.0000
	Philippine Peso	1.0000 ± 0.0000		Singapore Dollar	1.0000 ± 0.0000
	Mexican Peso	1.0000 ± 0.0000		Jamaican Dollar	1.0000 ± 0.0000
	Namibian Dollar	1.0000 ± 0.0000		Japanese Yen	1.0000 ± 0.0000
	Japanese Yen	1.0000 ± 0.0000		Cayman Islands Dollar	1.0000 ± 0.0000
Bottom 5	Mongolian Tögrög	0.0103 ± 0.0035	Bottom 5	Ghanaian Cedi	0.0142 ± 0.0066
	Tongan Paanga	0.0103 ± 0.0035		Bangladeshi Taka	0.0145 ± 0.0060
	Cambodian Riel	0.0105 ± 0.0036		Cambodian Riel	0.0163 ± 0.0071
	Azerbaijani Manat	0.0106 ± 0.0037		Costa Rican Colón	0.0168 ± 0.0077
	Lao Kip	0.0109 ± 0.0038		Czech Koruna	0.0173 ± 0.0044

Table 16: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: Alibaba-NLP gte-multilingual-base

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	Australian Dollar	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		Japanese Yen	1.0000 ± 0.0000
	British Pound Sterling	1.0000 ± 0.0000		Hong Kong Dollar	1.0000 ± 0.0000
	Euro	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	Danish Krone	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
Bottom 5	Macanese Pataca	0.0055 ± 0.0020	Bottom 5	Cape Verdean Escudo	0.0052 ± 0.0018
	North Korean Won	0.0057 ± 0.0021		Ghanaian Cedi	0.0087 ± 0.0040
	Bulgarian Lev	0.0064 ± 0.0026		Macanese Pataca	0.0090 ± 0.0042
	Venezuelan Bolívar	0.0064 ± 0.0026		Bosnia and Herzegovina Convertible Mark	0.0104 ± 0.0039
	Moroccan Dirham	0.0067 ± 0.0028		United Arab Emirates Dirham	0.0139 ± 0.0057
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Turkish Lira	1.0000 ± 0.0000	Top 5	Brazilian Real	1.0000 ± 0.0000
	Philippine Peso	1.0000 ± 0.0000		Bangladeshi Taka	1.0000 ± 0.0000
	Mexican Peso	1.0000 ± 0.0000		Japanese Yen	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
	Japanese Yen	1.0000 ± 0.0000		Indian Rupee	1.0000 ± 0.0000
Bottom 5	Armenian Dram	0.0108 ± 0.0038	Bottom 5	Cayman Islands Dollar	0.0161 ± 0.0039
	Azerbaijani Manat	0.0109 ± 0.0038		Sudanese Pound	0.0168 ± 0.0049
	Macanese Pataca	0.0110 ± 0.0039		Qatari Riyal	0.0188 ± 0.0051
	Afghan Afghani	0.0113 ± 0.0041		Armenian Dram	0.0219 ± 0.0093
	Mongolian Tögrög	0.0117 ± 0.0043		Solomon Islands Dollar	0.0231 ± 0.0063

Table 17: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: jina-embeddings-v2-base-en

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	1.0000 ± 0.0000	Top 5	Australian Dollar	1.0000 ± 0.0000
	Euro	0.7500 ± 0.2500		United States Dollar	1.0000 ± 0.0000
	British Pound Sterling	0.7500 ± 0.2500		Thai Baht	1.0000 ± 0.0000
	New Zealand Dollar	0.7500 ± 0.2500		New Zealand Dollar	1.0000 ± 0.0000
	Russian Ruble	0.7500 ± 0.2500		Japanese Yen	1.0000 ± 0.0000
Bottom 5	Mongolian Tögrög	0.0049 ± 0.0016	Bottom 5	Macanese Pataca	0.0061 ± 0.0022
	Tongan Paanga	0.0050 ± 0.0017		Kyrgyzstani Som	0.0063 ± 0.0021
	Eswatini Lilangeni	0.0052 ± 0.0018		Mongolian Tögrög	0.0064 ± 0.0019
	Lesotho Loti	0.0052 ± 0.0018		Chilean Peso	0.0065 ± 0.0011
	North Korean Won	0.0054 ± 0.0020		Ugandan Shilling	0.0066 ± 0.0012
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	1.0000 ± 0.0000	Top 5	Indian Rupee	1.0000 ± 0.0000
	Euro	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		New Zealand Dollar	1.0000 ± 0.0000
	British Pound Sterling	0.6666 ± 0.3334		Polish Złoty	1.0000 ± 0.0000
	Philippine Peso	0.5250 ± 0.4750		Trinidad and Tobago Dollar	1.0000 ± 0.0000
Bottom 5	Tongan Paanga	0.0101 ± 0.0034	Bottom 5	Macanese Pataca	0.0113 ± 0.0039
	Mongolian Tögrög	0.0103 ± 0.0035		Tongan Paanga	0.0115 ± 0.0040
	Paraguayan Guaraní	0.0105 ± 0.0036		Brazilian Real	0.0157 ± 0.0051
	Lao Kip	0.0106 ± 0.0037		Lebanese Pound	0.0182 ± 0.0030
	Cambodian Riel	0.0108 ± 0.0038		Solomon Islands Dollar	0.0188 ± 0.0039

Table 18: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: BAAI bge-m3

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	British Pound Sterling	0.6250 ± 0.3750	Top 5	Vanuatu Vatu	1.0000 ± 0.0000
	United States Dollar	0.6000 ± 0.4000		Ukrainian Hryvnia	1.0000 ± 0.0000
	Euro	0.5715 ± 0.4285		Lao Kip	1.0000 ± 0.0000
	Australian Dollar	0.5454 ± 0.4546		Lesotho Loti	1.0000 ± 0.0000
	Hong Kong Dollar	0.5416 ± 0.4584		Ethiopian Birr	1.0000 ± 0.0000
Bottom 5	Peruvian Sol	0.0050 ± 0.0017	Bottom 5	Moroccan Dirham	0.0052 ± 0.0015
	Macanese Pataca	0.0057 ± 0.0021		Albanian Lek	0.0083 ± 0.0025
	São Tomé and Príncipe Dobra	0.0057 ± 0.0022		Sudanese Pound	0.0100 ± 0.0035
	Georgian Lari	0.0060 ± 0.0023		Cuban Peso	0.0113 ± 0.0047
	Tajikistani Somoni	0.0066 ± 0.0027		São Tomé and Príncipe Dobra	0.0127 ± 0.0042
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	0.7500 ± 0.2500	Top 5	Belize Dollar	1.0000 ± 0.0000
	Euro	0.5625 ± 0.4375		Euro	1.0000 ± 0.0000
	New Zealand Dollar	0.5500 ± 0.4500		Russian Ruble	1.0000 ± 0.0000
	Fijian Dollar	0.5333 ± 0.4667		Polish Złoty	1.0000 ± 0.0000
	Turkish Lira	0.5200 ± 0.4800		Vietnamese Dong	1.0000 ± 0.0000
Bottom 5	Peruvian Sol	0.0101 ± 0.0034	Bottom 5	Kuwaiti Dinar	0.0167 ± 0.0055
	Armenian Dram	0.0106 ± 0.0037		Macedonian Denar	0.0213 ± 0.0081
	Macanese Pataca	0.0109 ± 0.0038		Peruvian Sol	0.0221 ± 0.0112
	Paraguayan Guaraní	0.0110 ± 0.0039		Paraguayan Guaraní	0.0226 ± 0.0097
	Georgian Lari	0.0112 ± 0.0040		Azerbaijani Manat	0.0277 ± 0.0069

Table 19: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: snowflake-arctic-embed-m-v1.5

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	Australian Dollar	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		United States Dollar	1.0000 ± 0.0000
	British Pound Sterling	0.6666 ± 0.3334		Japanese Yen	1.0000 ± 0.0000
	Japanese Yen	0.6666 ± 0.3334		Euro	1.0000 ± 0.0000
	United States Dollar	0.6666 ± 0.3334		British Pound Sterling	1.0000 ± 0.0000
Bottom 5	Costa Rican Colón	0.0050 ± 0.0017	Bottom 5	North Korean Won	0.0052 ± 0.0016
	Haitian Gourde	0.0050 ± 0.0017		Mauritanian Ouguiya	0.0052 ± 0.0017
	North Korean Won	0.0050 ± 0.0017		Cape Verdean Escudo	0.0054 ± 0.0019
	Bosnia and Herzegovina Convertible Mark	0.0052 ± 0.0018		Tajikistani Somoni	0.0056 ± 0.0019
	Nicaraguan Córdoba	0.0052 ± 0.0018		Bahamian Dollar	0.0057 ± 0.0009
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Indian Rupee	1.0000 ± 0.0000	Top 5	Euro	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		Chinese Yuan	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		Indian Rupee	1.0000 ± 0.0000
	Philippine Peso	0.5167 ± 0.4834		Polish Złoty	1.0000 ± 0.0000
	Mexican Peso	0.5084 ± 0.4915		New Zealand Dollar	1.0000 ± 0.0000
Bottom 5	Costa Rican Colón	0.0101 ± 0.0034	Bottom 5	Sudanese Pound	0.0107 ± 0.0031
	Panamanian Balboa	0.0103 ± 0.0035		Peruvian Sol	0.0113 ± 0.0036
	Israeli New Shekel	0.0105 ± 0.0036		Czech Koruna	0.0116 ± 0.0033
	Peruvian Sol	0.0108 ± 0.0038		Dominican Peso	0.0136 ± 0.0036
	Czech Koruna	0.0113 ± 0.0041		Liberian Dollar	0.0144 ± 0.0038

Table 20: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: sentence-transformers paraphrase-multilingual-mpnet-base-v2

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Euro	1.0000 ± 0.0000	Top 5	British Pound Sterling	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		New Zealand Dollar	1.0000 ± 0.0000
	United States Dollar	1.0000 ± 0.0000		United States Dollar	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	British Pound Sterling	1.0000 ± 0.0000		Swiss Franc	1.0000 ± 0.0000
Bottom 5	Eritrean Nakfa	0.0053 ± 0.0019	Bottom 5	Solomon Islands Dollar	0.0055 ± 0.0012
	Lesotho Loti	0.0053 ± 0.0018		Ugandan Shilling	0.0055 ± 0.0016
	Tongan Paanga	0.0053 ± 0.0018		Sudanese Pound	0.0056 ± 0.0014
	Salvadoran Colón	0.0055 ± 0.0020		Ghanaian Cedi	0.0057 ± 0.0019
	Afghan Afghani	0.0056 ± 0.0020		Bahamian Dollar	0.0058 ± 0.0009
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	New Zealand Dollar	1.0000 ± 0.0000	Top 5	Indian Rupee	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	United States Dollar	1.0000 ± 0.0000		Japanese Yen	1.0000 ± 0.0000
	Euro	0.7500 ± 0.2500		Hong Kong Dollar	1.0000 ± 0.0000
	British Pound Sterling	0.7500 ± 0.2500		New Zealand Dollar	1.0000 ± 0.0000
Bottom 5	Georgian Lari	0.0103 ± 0.0035	Bottom 5	Dominican Peso	0.0121 ± 0.0038
	Costa Rican Colón	0.0105 ± 0.0036		Uruguayan Peso	0.0138 ± 0.0041
	Israeli New Shekel	0.0106 ± 0.0037		Solomon Islands Dollar	0.0140 ± 0.0019
	Afghan Afghani	0.0109 ± 0.0038		Mongolian Tögrög	0.0142 ± 0.0042
	Paraguayan Guaraní	0.0112 ± 0.0040		Fijian Dollar	0.0151 ± 0.0021

Table 21: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: sentence-transformers all-mpnet-base-v2

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	Australian Dollar	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		New Zealand Dollar	1.0000 ± 0.0000
	United States Dollar	1.0000 ± 0.0000		United States Dollar	1.0000 ± 0.0000
	Japanese Yen	1.0000 ± 0.0000		Euro	1.0000 ± 0.0000
	British Pound Sterling	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
Bottom 5	Eswatini Lilangeni	0.0051 ± 0.0017	Bottom 5	Sudanese Pound	0.0052 ± 0.0014
	Paraguayan Guaraní	0.0052 ± 0.0018		Albanian Lek	0.0053 ± 0.0018
	Tongan Paanga	0.0052 ± 0.0018		Turkish Lira	0.0058 ± 0.0016
	Peruvian Sol	0.0053 ± 0.0019		Lesotho Loti	0.0059 ± 0.0022
	Mongolian Tögrög	0.0053 ± 0.0018		Ghanaian Cedi	0.0059 ± 0.0022
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	1.0000 ± 0.0000	Top 5	Bulgarian Lev	1.0000 ± 0.0000
	Chinese Yuan	1.0000 ± 0.0000		United States Dollar	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		New Zealand Dollar	1.0000 ± 0.0000
	Philippine Peso	1.0000 ± 0.0000		Philippine Peso	1.0000 ± 0.0000
	Polish Złoty	1.0000 ± 0.0000		Japanese Yen	1.0000 ± 0.0000
Bottom 5	Georgian Lari	0.0103 ± 0.0035	Bottom 5	Paraguayan Guaraní	0.0150 ± 0.0058
	Paraguayan Guaraní	0.0104 ± 0.0035		Costa Rican Colón	0.0152 ± 0.0071
	Costa Rican Colón	0.0105 ± 0.0036		Nigerian Naira	0.0154 ± 0.0059
	Tongan Paanga	0.0105 ± 0.0036		Georgian Lari	0.0162 ± 0.0061
	Afghan Afghani	0.0106 ± 0.0037		Armenian Dram	0.0166 ± 0.0062

Table 22: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: Qwen3-Embedding-0.6B

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	Bangladeshi Taka	1.0000 ± 0.0000
	Brazilian Real	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	Russian Ruble	1.0000 ± 0.0000		Sri Lankan Rupee	1.0000 ± 0.0000
	Japanese Yen	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
	South Korean Won	1.0000 ± 0.0000		Danish Krone	1.0000 ± 0.0000
Bottom 5	Armenian Dram	0.0050 ± 0.0017	Bottom 5	Seychellois Rupee	0.0055 ± 0.0017
	Comorian Franc	0.0050 ± 0.0017		Cayman Islands Dollar	0.0056 ± 0.0015
	Samoan Tālā	0.0052 ± 0.0018		Djiboutian Franc	0.0056 ± 0.0018
	Tongan Paanga	0.0052 ± 0.0018		Comorian Franc	0.0062 ± 0.0023
	Paraguayan Guaraní	0.0053 ± 0.0019		Samoan Tālā	0.0065 ± 0.0025
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Indian Rupee	1.0000 ± 0.0000	Top 5	Bulgarian Lev	1.0000 ± 0.0000
	Czech Koruna	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	Thai Baht	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
	British Pound Sterling	0.7500 ± 0.2500		Thai Baht	1.0000 ± 0.0000
	Mexican Peso	0.7500 ± 0.2500		Polish Złoty	1.0000 ± 0.0000
Bottom 5	Paraguayan Guaraní	0.0104 ± 0.0035	Bottom 5	Sudanese Pound	0.0112 ± 0.0034
	Armenian Dram	0.0105 ± 0.0036		Paraguayan Guaraní	0.0115 ± 0.0041
	Mongolian Tögrög	0.0106 ± 0.0037		Lao Kip	0.0135 ± 0.0054
	Lao Kip	0.0109 ± 0.0038		Surinamese Dollar	0.0182 ± 0.0062
	Syrian Pound	0.0110 ± 0.0039		Serbian Dinar	0.0186 ± 0.0077

Table 23: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: google embeddinggemma-300M

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	United Arab Emirates Dirham	1.0000 ± 0.0000
	Brazilian Real	1.0000 ± 0.0000		Australian Dollar	1.0000 ± 0.0000
	Danish Krone	1.0000 ± 0.0000		Botswana Pula	1.0000 ± 0.0000
	Czech Koruna	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	United States Dollar	1.0000 ± 0.0000		Thai Baht	1.0000 ± 0.0000
Bottom 5	Salvadoran Colón	0.0055 ± 0.0020	Bottom 5	Albanian Lek	0.0073 ± 0.0026
	Macanese Pataca	0.0056 ± 0.0021		Burundian Franc	0.0091 ± 0.0027
	Armenian Dram	0.0056 ± 0.0020		Somali Shilling	0.0095 ± 0.0035
	Angolan Kwanza	0.0057 ± 0.0021		South Sudanese Pound	0.0101 ± 0.0034
	Maldivian Rufiyaa	0.0060 ± 0.0023		Samoan Tālā	0.0122 ± 0.0061
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	1.0000 ± 0.0000	Top 5	United Arab Emirates Dirham	1.0000 ± 0.0000
	Chinese Yuan	1.0000 ± 0.0000		Bangladeshi Taka	1.0000 ± 0.0000
	Turkish Lira	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	Mexican Peso	1.0000 ± 0.0000		Bulgarian Lev	1.0000 ± 0.0000
	Philippine Peso	1.0000 ± 0.0000		Israeli New Shekel	1.0000 ± 0.0000
Bottom 5	Azerbaijani Manat	0.0101 ± 0.0034	Bottom 5	Lao Kip	0.0201 ± 0.0056
	Armenian Dram	0.0104 ± 0.0035		Syrian Pound	0.0218 ± 0.0085
	Lao Kip	0.0112 ± 0.0040		Sudanese Pound	0.0267 ± 0.0118
	Macanese Pataca	0.0115 ± 0.0041		Georgian Lari	0.0298 ± 0.0137
	Cambodian Riel	0.0118 ± 0.0043		Macanese Pataca	0.0304 ± 0.0150

Table 24: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: intfloat multilingual-e5-large-instruct

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Australian Dollar	1.0000 ± 0.0000	Top 5	Australian Dollar	1.0000 ± 0.0000
	Hong Kong Dollar	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	British Pound Sterling	1.0000 ± 0.0000		Hong Kong Dollar	1.0000 ± 0.0000
	Czech Koruna	1.0000 ± 0.0000		Czech Koruna	1.0000 ± 0.0000
	Danish Krone	1.0000 ± 0.0000		Danish Krone	1.0000 ± 0.0000
Bottom 5	Mongolian Tögrög	0.0050 ± 0.0017	Bottom 5	Cape Verdean Escudo	0.0057 ± 0.0019
	Tongan Paanga	0.0050 ± 0.0017		Sudanese Pound	0.0063 ± 0.0015
	Samoan Tālā	0.0052 ± 0.0018		Moroccan Dirham	0.0077 ± 0.0023
	Cape Verdean Escudo	0.0052 ± 0.0018		Albanian Lek	0.0077 ± 0.0026
	Papua New Guinean Kina	0.0054 ± 0.0020		Tongan Paanga	0.0082 ± 0.0029
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Hong Kong Dollar	1.0000 ± 0.0000	Top 5	Bulgarian Lev	1.0000 ± 0.0000
	Czech Koruna	1.0000 ± 0.0000		Bangladeshi Taka	1.0000 ± 0.0000
	Mexican Peso	1.0000 ± 0.0000		Brazilian Real	1.0000 ± 0.0000
	New Zealand Dollar	1.0000 ± 0.0000		British Pound Sterling	1.0000 ± 0.0000
	Polish Złoty	1.0000 ± 0.0000		Hong Kong Dollar	1.0000 ± 0.0000
Bottom 5	Paraguayan Guaraní	0.0101 ± 0.0034	Bottom 5	Serbian Dinar	0.0105 ± 0.0030
	Afghan Afghani	0.0103 ± 0.0035		Armenian Dram	0.0116 ± 0.0040
	Lao Kip	0.0105 ± 0.0036		Afghan Afghani	0.0125 ± 0.0044
	Tongan Paanga	0.0110 ± 0.0039		Lao Kip	0.0126 ± 0.0046
	Panamanian Balboa	0.0113 ± 0.0041		Ghanaian Cedi	0.0131 ± 0.0045

Table 25: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: intfloat multilingual-e5-large

Acronym → Currency Name			Currency Name → Acronym		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Norwegian Krone	0.5312 ± 0.4688	Top 5	Zimbabwean Gold	1.0000 ± 0.0000
	Danish Krone	0.5294 ± 0.4706		Vietnamese Dong	1.0000 ± 0.0000
	Icelandic Króna	0.5156 ± 0.4844		Thai Baht	1.0000 ± 0.0000
	Japanese Yen	0.5147 ± 0.4853		Polish Złoty	1.0000 ± 0.0000
	Swedish Krona	0.5147 ± 0.4853		Afghan Afghani	0.7500 ± 0.2500
Bottom 5	Tongan Paanga	0.0051 ± 0.0017	Bottom 5	Cuban Peso	0.0096 ± 0.0031
	Mongolian Tögrög	0.0052 ± 0.0018		Bolivian Boliviano	0.0103 ± 0.0032
	Lao Kip	0.0052 ± 0.0018		Moroccan Dirham	0.0115 ± 0.0049
	Samoan Tālā	0.0052 ± 0.0018		Sudanese Pound	0.0136 ± 0.0053
	Bolivian Boliviano	0.0057 ± 0.0021		Peruvian Sol	0.0141 ± 0.0048
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Czech Koruna	0.5161 ± 0.4839	Top 5	Bulgarian Lev	1.0000 ± 0.0000
	Mexican Peso	0.5147 ± 0.4853		Czech Koruna	1.0000 ± 0.0000
	Guyanese Dollar	0.5139 ± 0.4861		Nigerian Naira	1.0000 ± 0.0000
	Polish Złoty	0.5128 ± 0.4872		Kazakhstani Tenge	1.0000 ± 0.0000
	Hong Kong Dollar	0.5122 ± 0.4878		Ukrainian Hryvnia	1.0000 ± 0.0000
Bottom 5	Lao Kip	0.0101 ± 0.0034	Bottom 5	Lao Kip	0.0101 ± 0.0034
	Afghan Afghani	0.0103 ± 0.0035		Armenian Dram	0.0109 ± 0.0037
	Armenian Dram	0.0104 ± 0.0035		Ghanaian Cedi	0.0111 ± 0.0028
	Panamanian Balboa	0.0108 ± 0.0038		Paraguayan Guaraní	0.0114 ± 0.0029
	Azerbaijani Manat	0.0110 ± 0.0039		Costa Rican Colón	0.0122 ± 0.0033

Table 26: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: sentence-transformers LaBSE

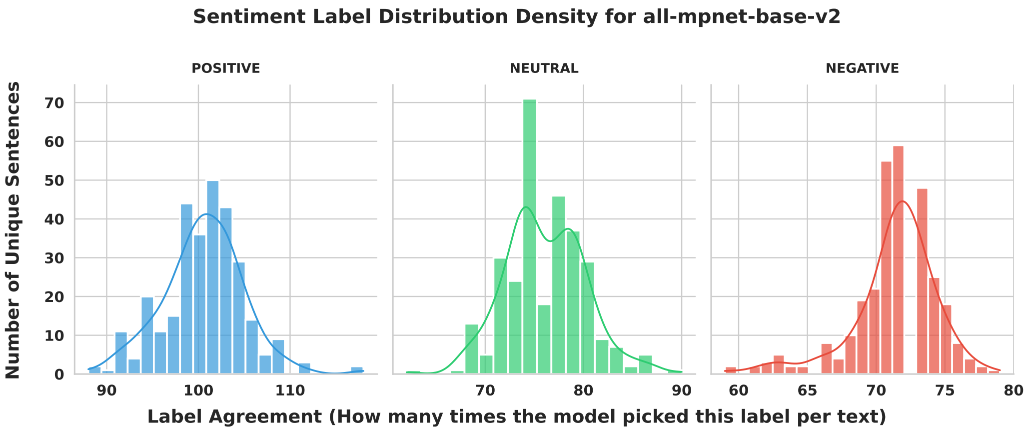
Acronym → Currency Name			**Currency Name → Acronym**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Japanese Yen	0.5094 ± 0.4905	Top 5	Brazilian Real	1.0000 ± 0.0000
	Vanuatu Vatu	0.5069 ± 0.4931		Japanese Yen	1.0000 ± 0.0000
	Thai Baht	0.5069 ± 0.4931		Thai Baht	1.0000 ± 0.0000
	United States Dollar	0.5054 ± 0.4945		Ukrainian Hryvnia	1.0000 ± 0.0000
	Lebanese Pound	0.5052 ± 0.4947		Kazakhstani Tenge	1.0000 ± 0.0000
Bottom 5	São Tomé and Príncipe Dobra	0.0059 ± 0.0022	Bottom 5	Albanian Lek	0.0101 ± 0.0044
	Bosnia and Herzegovina Convertible Mark	0.0061 ± 0.0024		East Caribbean Dollar	0.0103 ± 0.0042
	Eswatini Lilangeni	0.0066 ± 0.0027		Samoan Tālā	0.0110 ± 0.0059
	Samoan Tālā	0.0067 ± 0.0028		Bosnia and Herzegovina Convertible Mark	0.0175 ± 0.0064
	Papua New Guinean Kina	0.0078 ± 0.0037		Belarusian Ruble	0.0231 ± 0.0138
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	United States Dollar	0.5122 ± 0.4878	Top 5	Polish Złoty	1.0000 ± 0.0000
	Hong Kong Dollar	0.5086 ± 0.4914		Euro	1.0000 ± 0.0000
	Euro	0.5084 ± 0.4915		Hong Kong Dollar	0.7500 ± 0.2500
	Belize Dollar	0.5080 ± 0.4920		Japanese Yen	0.7500 ± 0.2500
	Fijian Dollar	0.5078 ± 0.4922		United States Dollar	0.7500 ± 0.2500
Bottom 5	Paraguayan Guaraní	0.0101 ± 0.0034	Bottom 5	Azerbaijani Manat	0.0102 ± 0.0033
	Afghan Afghani	0.0115 ± 0.0041		Afghan Afghani	0.0103 ± 0.0035
	Israeli New Shekel	0.0126 ± 0.0049		Georgian Lari	0.0103 ± 0.0035
	Ghanaian Cedi	0.0134 ± 0.0055		Paraguayan Guaraní	0.0106 ± 0.0037
	Armenian Dram	0.0134 ± 0.0055		Peruvian Sol	0.0178 ± 0.0072

Table 27: Syntax Gap: Summary of Top/Bottom 5 Currencies by Mean ± Std MRR Across Scenarios for Model: ProsusAI finbert

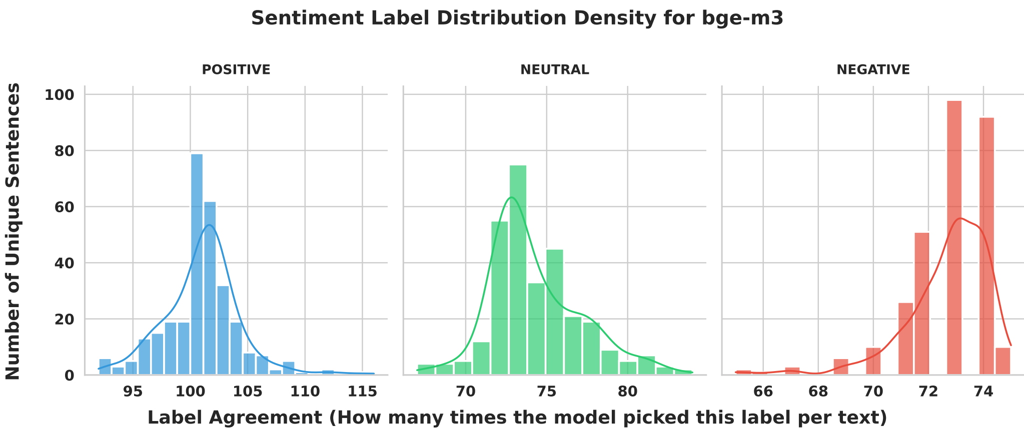
Acronym → Currency Name			**Currency Name → Acronym**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Japanese Yen	0.5063 ± 0.4936	Top 5	Mexican Peso	0.5263 ± 0.4737
	Armenian Dram	0.5055 ± 0.4945		Swedish Krona	0.5250 ± 0.4750
	Serbian Dinar	0.2537 ± 0.2463		Peruvian Sol	0.5084 ± 0.4915
	Azerbaijani Manat	0.2536 ± 0.2464		Armenian Dram	0.2833 ± 0.2167
	Swedish Krona	0.1705 ± 0.1628		United States Dollar	0.2708 ± 0.2291
Bottom 5	Somali Shilling	0.0049 ± 0.0016	Bottom 5	Eswatini Lilangeni	0.0051 ± 0.0017
	Kenyan Shilling	0.0050 ± 0.0017		Ugandan Shilling	0.0051 ± 0.0017
	Ugandan Shilling	0.0050 ± 0.0017		Bangladeshi Taka	0.0052 ± 0.0016
	South Sudanese Pound	0.0051 ± 0.0017		Tanzanian Shilling	0.0052 ± 0.0018
	Tanzanian Shilling	0.0051 ± 0.0017		Eritrean Nakfa	0.0053 ± 0.0018
Symbol → Currency Name			**Currency Name → Symbol**		
Type	**Currency**	**Mean ± Std MRR**	**Type**	**Currency**	**Mean ± Std MRR**
Top 5	Japanese Yen	0.5555 ± 0.4445	Top 5	Mexican Peso	0.6250 ± 0.3750
	Bulgarian Lev	0.5109 ± 0.4892		Bulgarian Lev	0.6000 ± 0.4000
	Mexican Peso	0.5107 ± 0.4894		British Pound Sterling	0.5357 ± 0.4643
	Belize Dollar	0.5098 ± 0.4902		Japanese Yen	0.5357 ± 0.4643
	Singapore Dollar	0.5094 ± 0.4905		Polish Złoty	0.1845 ± 0.1488
Bottom 5	Somali Shilling	0.0101 ± 0.0034	Bottom 5	Tunisian Dinar	0.0103 ± 0.0034
	Sudanese Pound	0.0103 ± 0.0035		Algerian Dinar	0.0104 ± 0.0035
	Afghan Afghani	0.0104 ± 0.0035		Libyan Dinar	0.0107 ± 0.0036
	United Arab Emirates Dirham	0.0106 ± 0.0037		Jordanian Dinar	0.0109 ± 0.0036
	Macanese Pataca	0.0106 ± 0.0037		Qatari Riyal	0.0109 ± 0.0038

E Sentiment Analysis Task

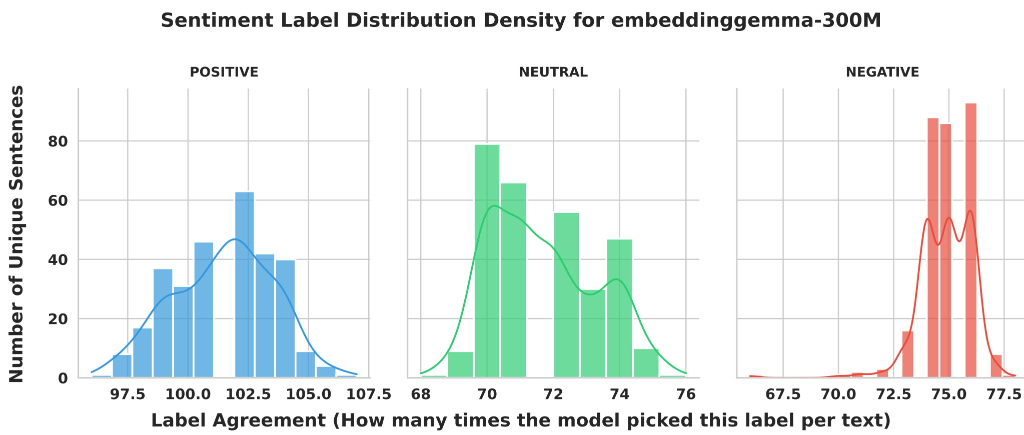
This section provides a detailed visualization of the performance and fairness of various text embedding models when subjected to currency-based perturbations in a sentiment analysis task. Figure 18 presents a Stability Probing analysis across multiple parts, tracking how sentiment predictions fluctuate across diverse currency probes for models ranging from general-purpose embeddings like all-mpnet-base-v2 and GTE to specialized financial models like FinBERT. These stability plots reveal the sensitivity of each model's output to specific denominations and regional identifiers. Complementing this, Figure 19 provides Currency-level Bias Plots, which quantify the systematic skew in sentiment classification. By aggregating results across models such as BGE-M3, Embedding Gemma, and Snowflake Arctic, these visualizations highlight the degree to which embedding models exhibit disparate sentiment associations based solely on the currency mentioned in the input text, thereby uncovering underlying currency-induced biases.



(a) Stability Probing: all-mpnet-base-v2

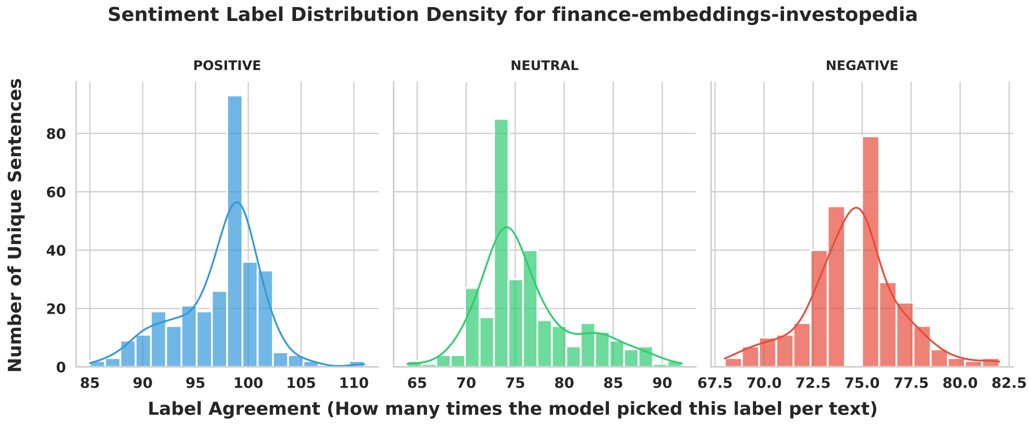


(b) Stability Probing: BGE-M3

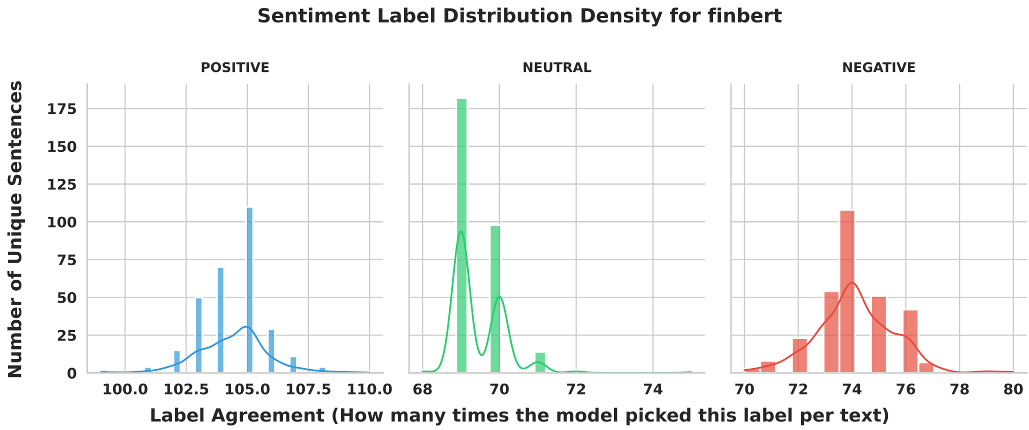


(c) Stability Probing: Embedding Gemma 300M

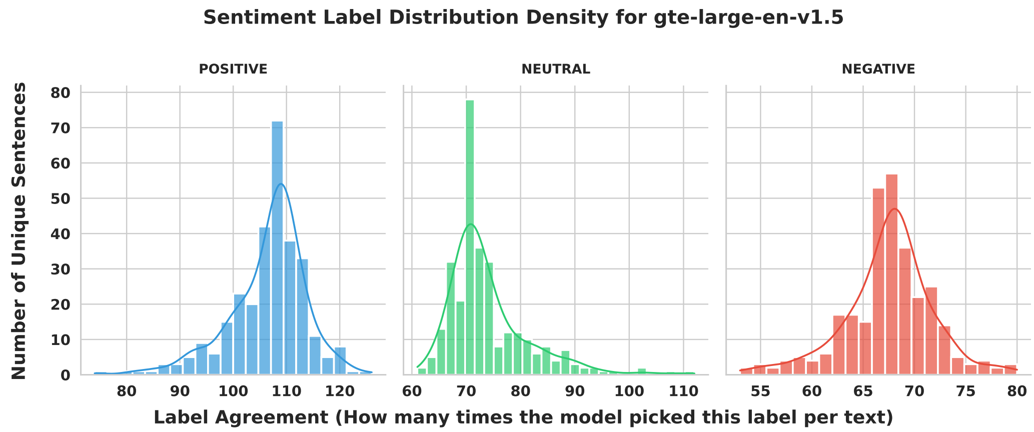
Figure 18: Stability of sentiment prediction across currency probes (Part I).



(d) Stability Probing: finance-embeddings-investopedia

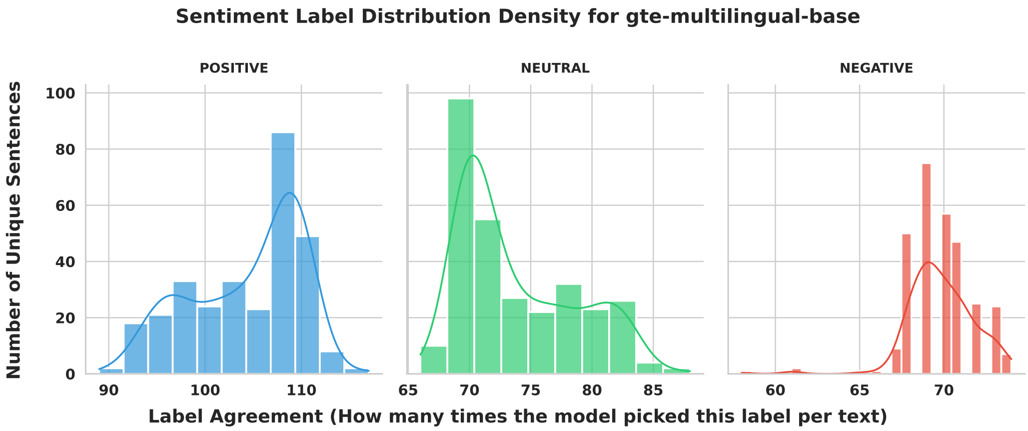


(e) Stability Probing: FinBERT

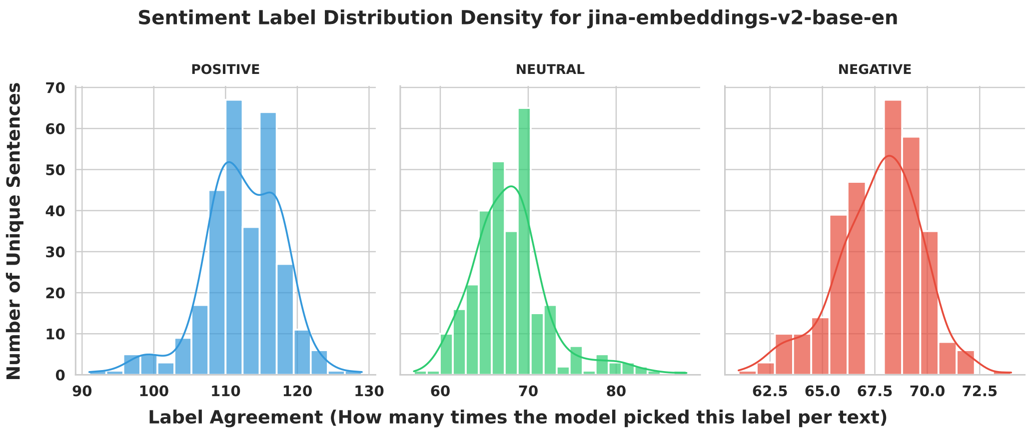


(f) Stability Probing: gte-large-en-v1.5

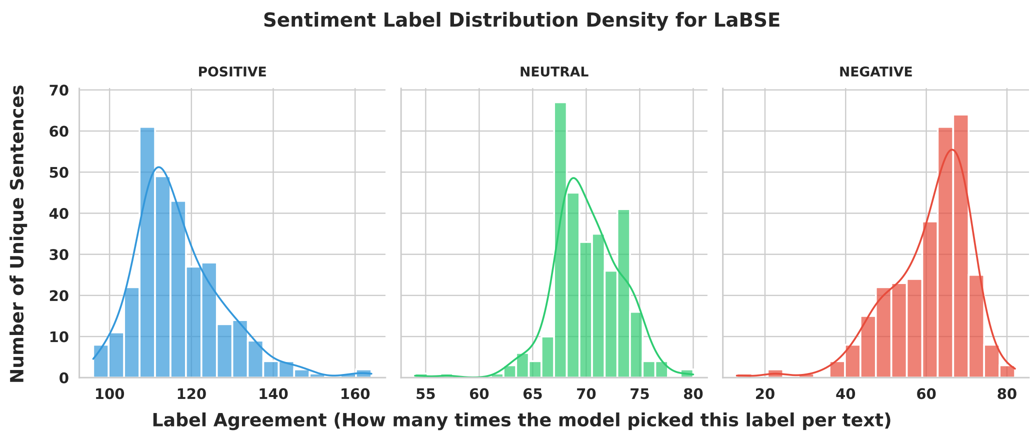
Figure 18: Stability of sentiment prediction across currency probes (Part II).



(g) Stability Probing: gte-multilingual-base

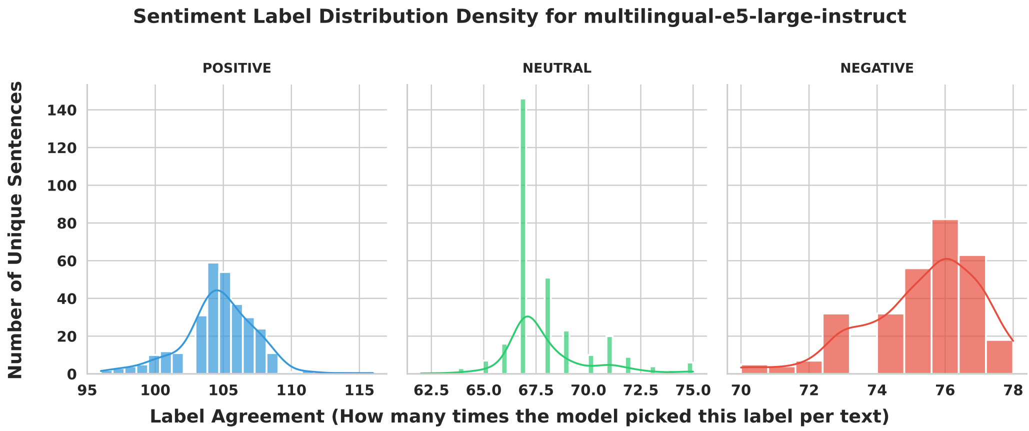


(h) Stability Probing: jina-embeddings-v2-base-en

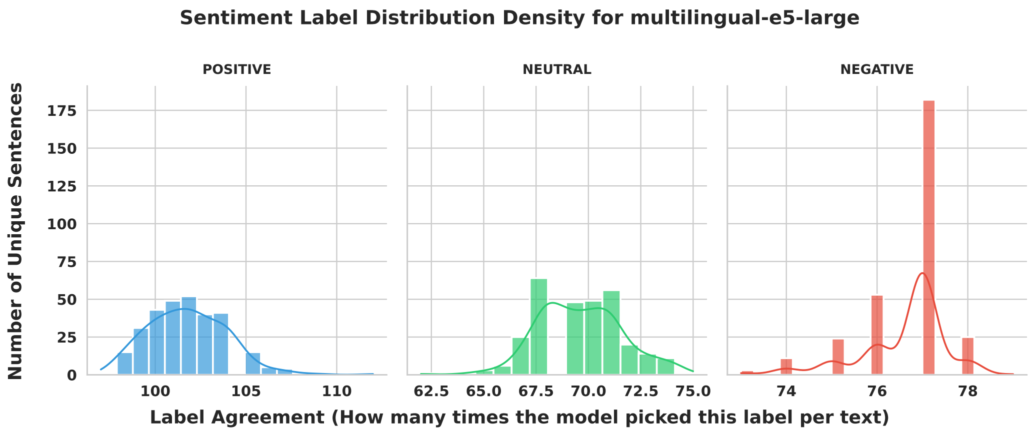


(i) Stability Probing: LaBSE

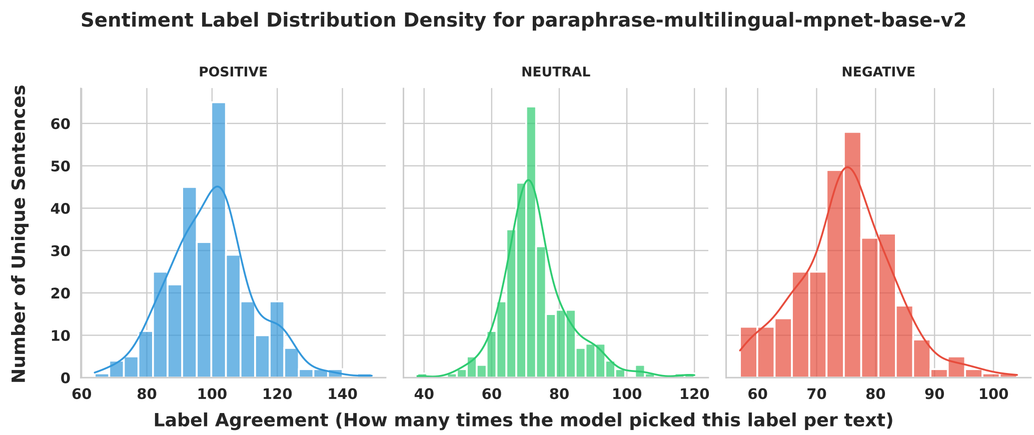
Figure 18: Stability of sentiment prediction across currency probes (Part III).



(j) Stability Probing: multilingual-e5-large-instruct

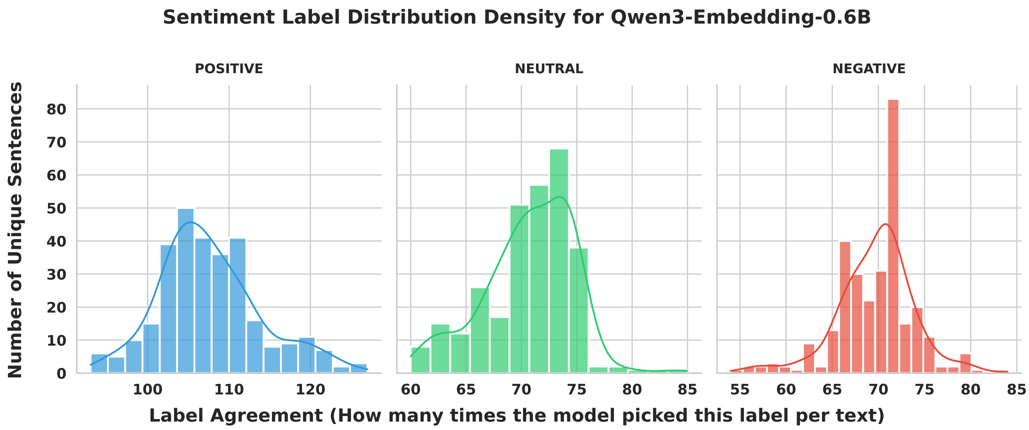


(k) Stability Probing: multilingual-e5-large

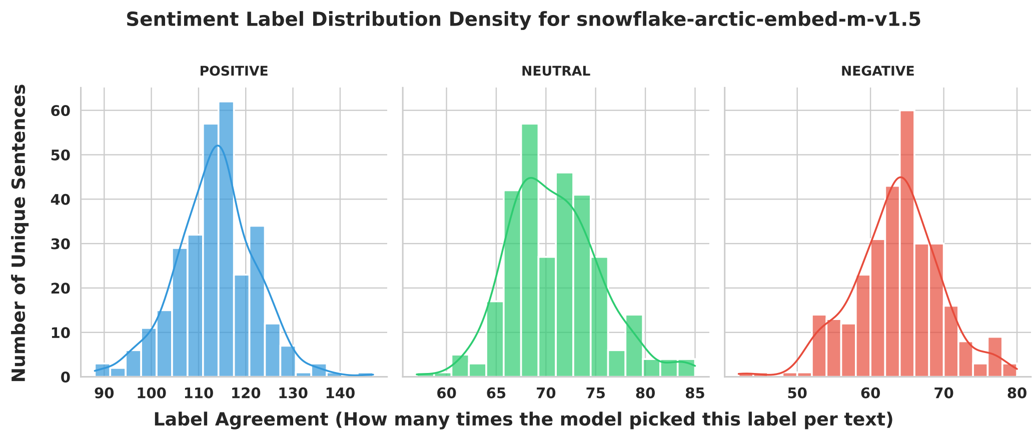


(l) Stability Probing: paraphrase-multilingual-mpnet-base-v2

Figure 18: Stability of sentiment prediction across currency probes (Part IV).

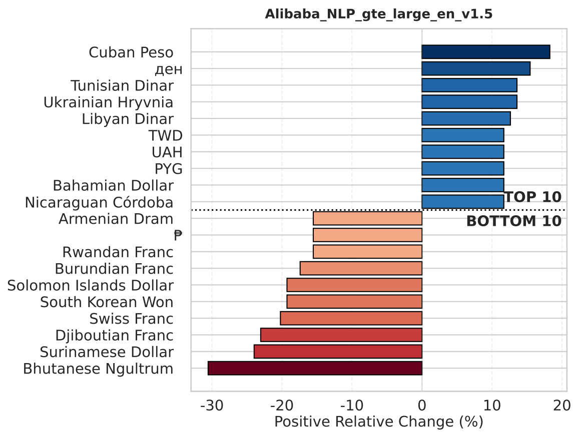


(m) Stability Probing: Qwen3-Embedding-0.6B

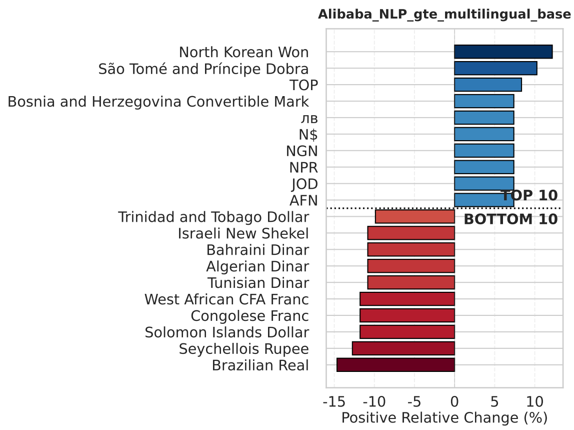


(n) Stability Probing: snowflake-arctic-embed-m-v1.5

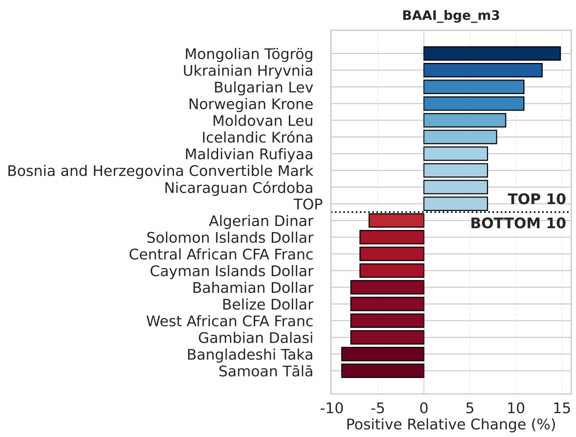
Figure 18: Stability of sentiment prediction across currency probes (Part V).



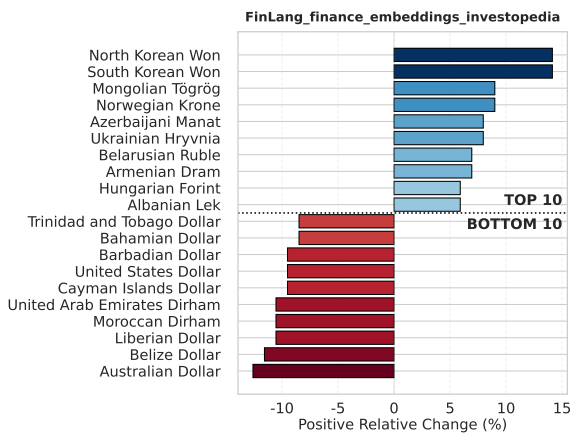
(a) Bias Plot:gte-large-en-v1.5



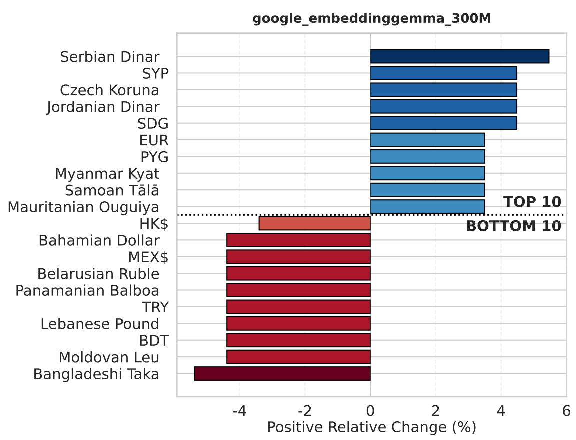
(b) Bias Plot: gte-multilingual-base



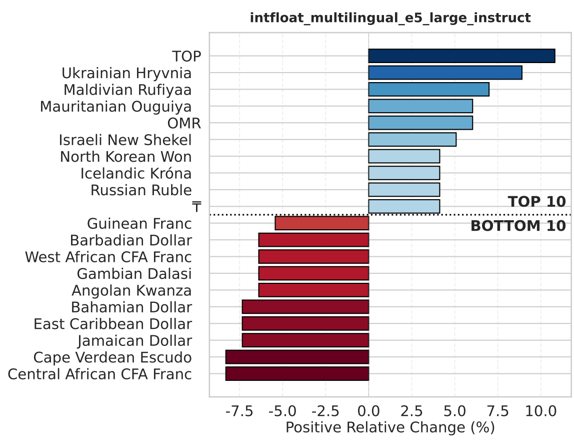
(c) Bias Plot: bge-m3



(d) Bias Plot: inance-embeddings-investopedia

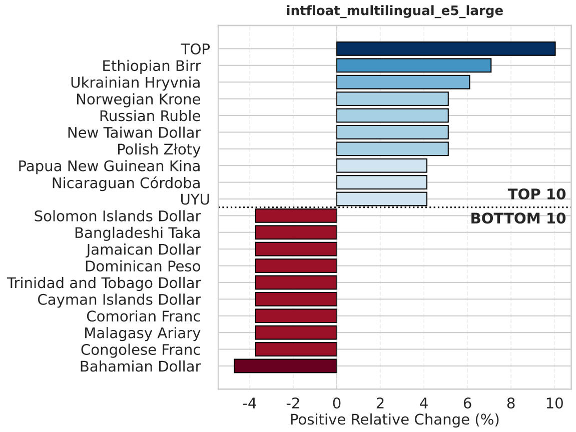


(e) Bias Plot: embeddinggemma-300M

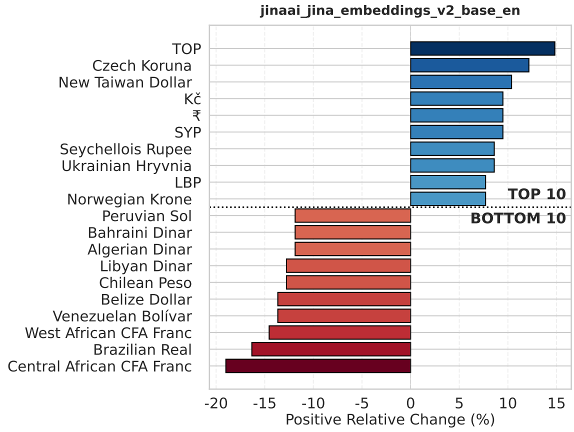


(f) Bias Plot: multilingual-e5-large-instruct

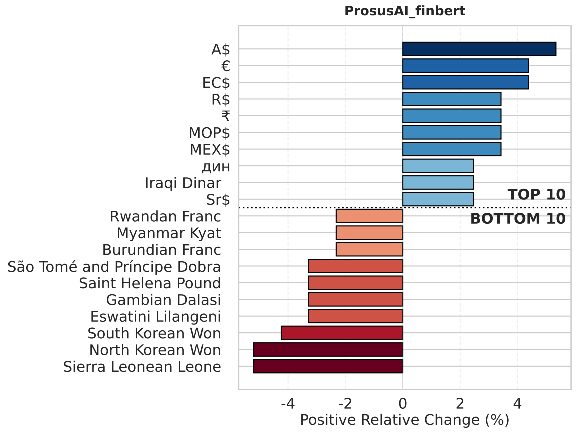
Figure 19: Currency-level bias plots for sentiment classification (Part I).



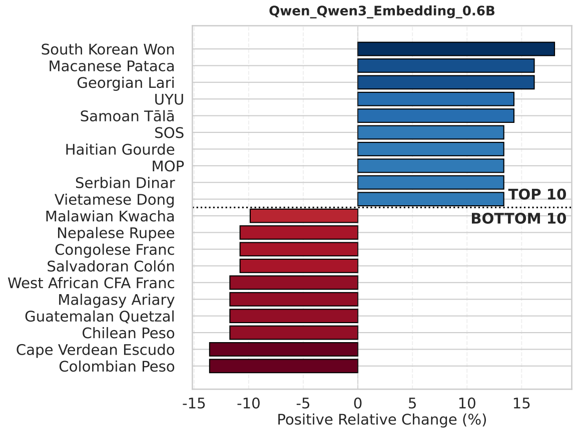
(g) Bias Plot: multilingual-e5-large



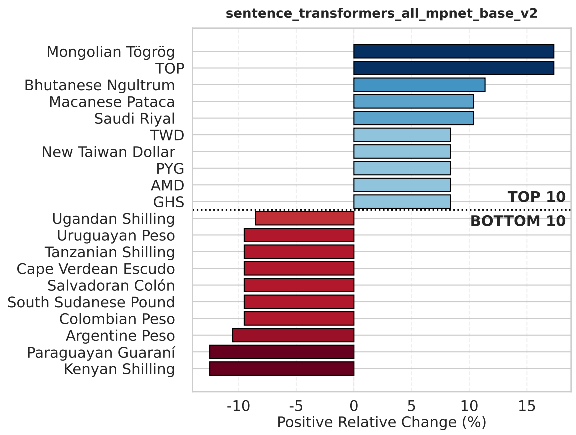
(h) Bias Plot: jina-embeddings-v2-base-en



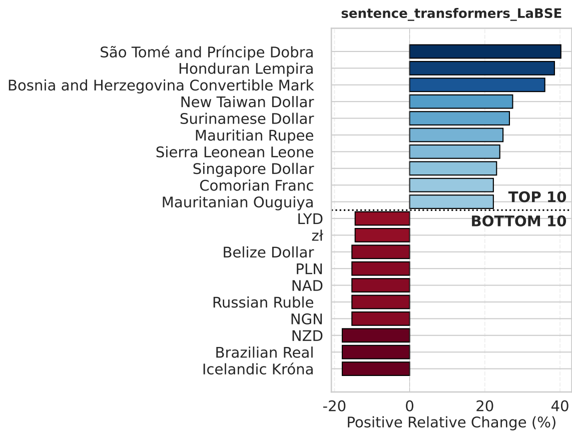
(i) Bias Plot: FinBERT



(j) Bias Plot: Qwen3-Embedding-0.6B

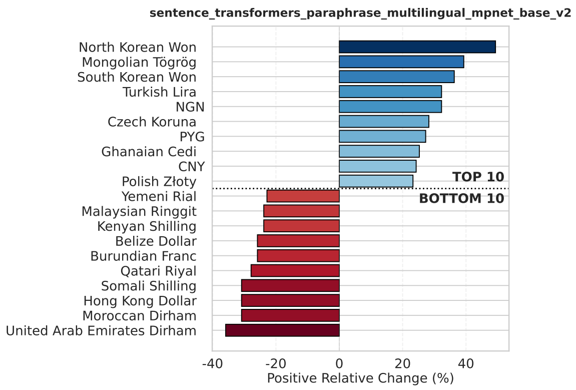


(k) Bias Plot: all-mpnet-base-v2

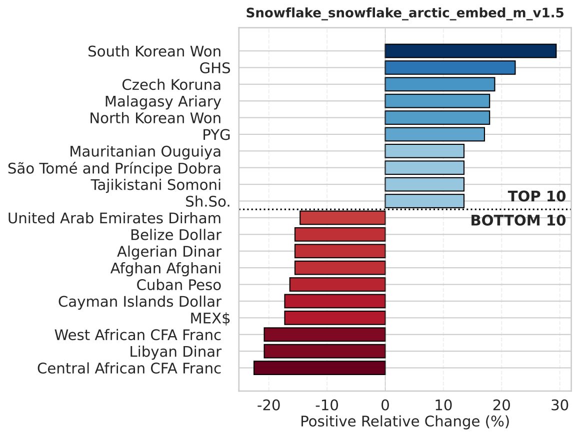


(l) Bias Plot: LaBSE

Figure 19: Currency-level bias plots for sentiment classification (Part II).



(m) Bias Plot: paraphrase-multilingual-mpnet-base-v2



(n) Bias Plot: snowflake-arctic-embed-m-v1.5

Figure 19: Currency-level bias plots for sentiment classification (Part III).

F Credit Default Prediction Task

The following tables report results for four representative embedding models under two settings: (i) notation-only perturbations and (ii) currency rate-adjusted perturbations. Across both settings, we observe consistent patterns of decision instability, including non-trivial flip rates and increased prediction entropy, indicating that the impact of currency variations is systematic and not model-specific. General-purpose multilingual models like *multilingual-e5-large* also exhibit high flip rates and prediction drift, changing credit decisions solely based on notation. Further, even when the underlying exchange rate is preserved, we still see behavior instability observed previously.

Table 28: **Credit Default Prediction: Performance across currency perturbations**: Decision flips and model stability. Exchange rate not applied.

Model	Flip Rate(%)	Pos. Rate (%)	AUCPR	Entropy
bge-m3	5.86 ± 2.63	29.40 ± 4.02	0.1781 ± 0.0072	0.1112 ± 0.2433
Qwen3-Embedding-0.6B	5.42 ± 3.56	34.39 ± 6.09	0.1685 ± 0.0030	0.1643 ± 0.2983
multilingual-e5-large	6.64 ± 7.09	42.77 ± 7.16	0.1591 ± 0.0020	0.1619 ± 0.2775
LaBSE	5.78 ± 3.98	29.36 ± 6.13	0.1586 ± 0.0070	0.1847 ± 0.2964

Table 29: **Credit Default Prediction: Performance across currency perturbations (currency rate adjusted)** leading to decision flips.

Model	Flip Rate (%)	Pos. Rate (%)	AUCPR	Entropy
bge-m3	8.06 ± 3.09	29.57 ± 4.92	0.1654 ± 0.0120	0.1718 ± 0.3068
Qwen3-Embedding-0.6B	7.50 ± 3.25	34.98 ± 6.38	0.1625 ± 0.0052	0.2012 ± 0.3295
multilingual-e5-large	7.79 ± 7.20	43.35 ± 7.65	0.1584 ± 0.0043	0.1868 ± 0.2959
LaBSE	15.10 ± 5.29	26.23 ± 10.38	0.1368 ± 0.0159	0.36 ± 0.3853

G Tokenization Statistics and Model Capability

We provide additional analysis on tokenization effects and their relationship with retrieval performance and associative bias.

Tokenization Statistics We analyze token-level statistics across different currency notation types—names, acronyms, and symbols—across multiple embedding models. Table 30 summarizes the average number of tokens and the occurrence of unknown (UNK/OOV) tokens.

Across all models, currency names and acronyms exhibit zero UNK occurrences, while certain currency symbols result in non-zero UNK counts. As expected, currency names have higher token lengths compared to acronyms and symbols.

Correlation with Retrieval Performance We examine whether tokenization artifacts influence a model's ability to retrieve correct currency representations, measured using Reciprocal Rank (RR). Table 33 reports Spearman correlations between RR and token-level features.

We observe weak and inconsistent correlations between token count and retrieval performance across models (with p-values $\gg 0.05$ in most cases). While some moderate correlations exist between UNK presence (for symbols) and RR, these effects are not consistent across models.

Correlation with Associative Bias We further analyze whether associative bias (measured via median rank differences) can be explained by tokenization artifacts. Across models and notation types, correlations between bias and token count are generally weak. Correlations with UNK presence are also limited and primarily observed in a few symbol-based cases.

These findings suggest that tokenization alone does not explain the observed biases, motivating our analysis across different currency representations (names, acronyms, and symbols).

Equal Token Length, Divergent Bias To illustrate that semantic associations—not tokenization—drive bias, we examine currencies with identical token lengths. For example, in *multilingual-e5-large*, "Omani Riyal," "Jordanian Dinar," and "Turkish Lira" each consist of four tokens. Despite this, they exhibit markedly different associative behaviors: Omani Riyal shows a positive median rank difference (+4.00, associated with luxury), while Jordanian Dinar (-4.50) and Turkish Lira (-4.00) are associated with poverty.

Overall, these results suggest that token length alone does not explain the observed syntax gap or associative biases. Notably, currencies with identical token counts can still exhibit significantly different bias profiles, indicating that semantic factors play a more dominant role. While moderate correlations are observed in a few cases—primarily within symbol representations due to UNK occurrences—these effects are sparse and not consistent across models.

Model Capability We observe that higher-capacity multilingual models (e.g.,multilingual e5-large, multilingual e5-large-instruct) in general exhibit lower syntax gaps and downstream volatility in comparison to others, suggesting that model scale augmented with multilingual training partially could be useful for reducing such biases. These findings suggest that these limitations are not confined to a specific architecture, but rather reflects representational issues across embedding models.

Table 30: **Tokenization statistics across currency notation types.** We report average number of tokens and UNK/OOV occurrences for acronyms, names, and symbols.

Model	Acronym		Name		Symbol	
	Avg. Tokens	UNK Count	Avg. Tokens	UNK Count	Avg. Tokens	UNK Count
gte-large-en-v1.5	1.97	0	3.55	0	2.04	19
gte-multilingual-base	1.89	0	3.97	0	2.27	7
bge-m3	1.89	0	3.97	0	2.27	7
finance-embeddings-investopedia	1.97	0	3.55	0	2.04	19
finbert	1.97	0	3.55	0	2.04	19
Qwen3-Embedding-0.6B	1.88	0	4.70	0	2.05	0
snowflake-arctic-embed-m-v1.5	1.97	0	3.55	0	2.04	19
embeddinggemma-300M	1.73	0	3.85	0	2.05	0
multilingual-e5-large	1.89	0	3.97	0	2.27	7
jina-embeddings-v2-base-en	1.97	0	3.55	0	2.04	19
LaBSE	1.52	0	3.12	0	1.92	4
all-mpnet-base-v2	1.97	0	3.55	0	2.04	19
paraphrase-multilingual-mpnet-base-v2	1.89	0	3.97	0	2.27	7

Table 31: **Correlation between tokenization features and associative bias (Luxury vs. Poverty).**

Model	Acronym		Name		Symbol	
	Tokens	UNK	Tokens	UNK	Tokens	UNK
gte-large-en-v1.5	0.1235	-	-0.0782	-	-0.4785	0.3365
gte-multilingual-base	-0.1150	-	-0.2027	-	-0.2323	0.1215
bge-m3	-0.3090	-	-0.1124	-	-0.0335	0.1597
finance-embeddings-investopedia	-0.0647	-	-0.2497	-	-0.2683	0.3367
finbert	0.0684	-	-0.0087	-	-0.3997	0.6543
Qwen3-Embedding-0.6B	-0.1418	-	-0.2389	-	0.4031	-
snowflake-arctic-embed-m-v1.5	-0.1839	-	-0.1004	-	-0.4952	0.6977
embeddinggemma-300M	-0.0007	-	0.1358	-	0.4401	-
multilingual-e5-large	-0.0761	-	-0.1937	-	0.3644	-0.2435
LaBSE	0.0209	-	-0.0229	-	0.0510	0.3932

Table 32: **Correlation between tokenization features and associative bias (Malice vs. Safety).**

Model	Acronym		Name		Symbol	
	Tokens	UNK	Tokens	UNK	Tokens	UNK
gte-large-en-v1.5	-0.0038	-	-0.1680	-	-0.7816	0.7683
gte-multilingual-base	-0.0417	-	-0.2304	-	-0.1070	0.3021
bge-m3	-0.2028	-	-0.0795	-	0.3483	0.3393
finance-embeddings-investopedia	-0.0195	-	0.1145	-	-0.3317	0.4185
finbert	0.0973	-	-0.1641	-	0.2094	-0.2738
Qwen3-Embedding-0.6B	0.1717	-	0.0732	-	-0.0868	-
snowflake-arctic-embed-m-v1.5	-0.1258	-	-0.2229	-	-0.1769	-0.0621
embeddinggemma-300M	0.0271	-	-0.3068	-	-0.2282	-
multilingual-e5-large	-0.0674	-	-0.0424	-	0.1573	0.4282
LaBSE	-0.2552	-	-0.1771	-	-0.1753	-0.3041

H Implementation Details

H.1 Model size

Table 34 depicts the different models and their size.

H.2 Packages Used

We utilized the `scikit-learn` (Pedregosa et al., 2011) package for computing statistical metrics, including cosine similarity and Mean Reciprocal Rank (MRR). For sentiment classification probing, we employed the `LogisticRegression` implementation from the same library. Data manipulation

Table 33: **Spearman correlation between tokenization features and Reciprocal Rank (RR).**

Model	**Acronym RR**		**Symbol RR**	
	Num Tokens	UNK Presence	Num Tokens	UNK Presence
gte-large-en-v1.5	0.0245	-	0.2017	-0.5645
gte-multilingual-base	-0.0721	-	-0.2886	-0.2890
bge-m3	-0.0847	-	-0.1591	-0.3023
finance-embeddings-investopedia	-0.0839	-	-0.0700	-0.3360
finbert	-0.0002	-	0.1091	-0.2149
Qwen3-Embedding-0.6B	0.0797	-	0.0048	-
snowflake-arctic-embed-m-v1.5	-0.1205	-	-0.0125	-0.2916
embeddinggemma-300M	-0.1103	-	-0.3088	-
multilingual-e5-large	-0.0429	-	-0.2801	-0.3883
LaBSE	-0.1905	-	-0.0656	-0.1973

Model Name	**Size (MB)**
finance-embeddings-investopedia	440
gte-large-en-v1.5	670
gte-multilingual-base	560
jina-embeddings-v2-base-en	273
bge-m3	1,140
snowflake-arctic-embed-m-v1.5	438
paraphrase-multilingual-mpnet-base-v2	1,110
all-mpnet-base-v2	438
Qwen3-Embedding-0.6B	1,200
embeddinggemma-300M	600
multilingual-e5-large-instruct	1,120
multilingual-e5-large	1,120
LaBSE	1,800
finbert	440

Table 34: The 14 text-embedding architectures evaluated in our benchmarks, representing a mix of domain-specific, general-purpose, and LLM-based embedding models.

and preprocessing were handled using `pandas` and `numpy`.

Our implementation relies on the `sentence-transformers` framework (Reimers, 2019), which is distributed under the Apache License, Version 2.0. All evaluated model weights were sourced from the Hugging Face Model Hub, governed by their respective open-source licenses (primarily Apache 2.0 and MIT). The artifacts and synthetic datasets generated for this research are available for non-commercial scientific use.

H.3 Model Information and Computational Budget

In Table 34, we present the specific configurations and storage footprints of the fourteen open-source embedding models evaluated in our study. This selection encompasses domain-specialized encoders (e.g., *FinBERT*, *finance-embeddings-investopedia*), general-purpose multilingual architectures (e.g., *LaBSE*, *multilingual-e5-large*), and recent LLM-based embedding models (e.g., *Qwen3-Embedding-0.6B*, *embeddinggemma-300M*). All experiments were conducted on a single 32GB GPU, with a total computational consumption of approximately 90 GPU hours for the full suite of retrieval and bias benchmarks.

H.4 Link to Code

Our code and evaluation framework are available at `https://github.com/sahilm1992/currency_bias_gap`.

I Currency Reference Table

The symbolic vocabulary was curated to isolate representational failures from simple lexical ambiguity. Ambiguous Single-Letter Tokens: Identifiers consisting of a single unadorned letter—such as "L" (Leone, Lempira) or "M"—were removed. These frequently appear in training corpora as units of measurement (length, mass). Symbols shared by multiple sovereign such as Won (₩), were omitted. Generic symbols, such as the lone "$", were systematically replaced with Contextualized Symbolic Identifiers (e.g., US$, MEX$, NZ$).

Currency (Code) [Symbol used?]	**Currency (Code) [Symbol used?]**	**Currency (Code) [Symbol used?]**
United States Dollar (USD) ✓	Bahraini Dinar (BHD) ✓	Moroccan Dirham (MAD) ✓
Indian Rupee (INR) ✓	Burundian Franc (BIF) ×	Moldovan Leu (MDL) ×
Euro (EUR) ✓	Bermudian Dollar (BMD) ×	Malagasy Ariary (MGA) ×
British Pound Sterling (GBP) ✓	Brunei Dollar (BND) ×	Macedonian Denar (MKD) ✓
Japanese Yen (JPY) ✓	Bolivian Boliviano (BOB) ×	Myanmar Kyat (MMK) ×
Swiss Franc (CHF) ×	Bahamian Dollar (BSD) ×	Mongolian Tögrög (MNT) ✓
North Korean Won (KPW) ×	Bhutanese Ngultrum (BTN) ×	Macanese Pataca (MOP) ✓
Pakistani Rupee (PKR) ×	Botswana Pula (BWP) ×	Mauritanian Ouguiya (MRU) ×
Chinese Yuan (CNY) ✓	Belarusian Ruble (BYN) ×	Mauritian Rupee (MUR) ×
Australian Dollar (AUD) ✓	Belize Dollar (BZD) ✓	Maldivian Rufiyaa (MVR) ×
Canadian Dollar (CAD) ✓	Congolese Franc (CDF) ×	Mozambican Metical (MZN) ×
Mexican Peso (MXN) ✓	Chilean Peso (CLP) ×	Namibian Dollar (NAD) ✓
Russian Ruble (RUB) ✓	Colombian Peso (COP) ×	Nicaraguan Córdoba (NIO) ×
South African Rand (ZAR) ×	Costa Rican Colón (CRC) ✓	Norwegian Krone (NOK) ×
Nigerian Naira (NGN) ✓	Cuban Peso (CUP) ×	New Zealand Dollar (NZD) ✓
Thai Baht (THB) ✓	Cape Verdean Escudo (CVE) ×	Panamanian Balboa (PAB) ✓
Vietnamese Dong (VND) ✓	Czech Koruna (CZK) ✓	Peruvian Sol (PEN) ✓
Indonesian Rupiah (IDR) ×	Djiboutian Franc (DJF) ×	Papua New Guinean Kina (PGK) ×
South Korean Won (KRW) ×	Danish Krone (DKK) ×	Paraguayan Guaraní (PYG) ✓
Polish Złoty (PLN) ✓	Dominican Peso (DOP) ✓	Qatari Riyal (QAR) ✓
Swedish Krona (SEK) ×	Eritrean Nakfa (ERN) ×	Romanian Leu (RON) ×
Turkish Lira (TRY) ✓	Ethiopian Birr (ETB) ×	Serbian Dinar (RSD) ✓
Brazilian Real (BRL) ✓	Fijian Dollar (FJD) ✓	Rwandan Franc (RWF) ×
United Arab Emirates Dirham (AED) ✓	Georgian Lari (GEL) ✓	Solomon Islands Dollar (SBD) ✓
Saudi Riyal (SAR) ✓	Gambian Dalasi (GMD) ×	Sudanese Pound (SDG) ✓
Seychellois Rupee (SCR) ×	Guinean Franc (GNF) ×	Saint Helena Pound (SHP) ×
Yemeni Rial (YER) ×	Guatemalan Quetzal (GTQ) ×	Somali Shilling (SOS) ✓
Sierra Leonean Leone (SLL) ×	Guyanese Dollar (GYD) ✓	Surinamese Dollar (SRD) ✓
Malawian Kwacha (MWK) ×	Hong Kong Dollar (HKD) ✓	South Sudanese Pound (SSP) ×
Ghanaian Cedi (GHS) ✓	Honduran Lempira (HNL) ×	São Tomé and Príncipe Dobra (STN) ×
Ugandan Shilling (UGX) ×	Haitian Gourde (HTG) ×	Salvadoran Colón (SVC) ×
Tanzanian Shilling (TZS) ×	Hungarian Forint (HUF) ×	Syrian Pound (SYP) ✓
West African CFA Franc (XOF) ×	Israeli New Shekel (ILS) ✓	Eswatini Lilangeni (SZL) ×
Algerian Dinar (DZD) ✓	Iraqi Dinar (IQD) ✓	Tajikistani Somoni (TJS) ×
Egyptian Pound (EGP) ✓	Iranian Rial (IRR) ×	Turkmenistani Manat (TMT) ×
Philippine Peso (PHP) ✓	Icelandic Króna (ISK) ×	Tunisian Dinar (TND) ✓
Malaysian Ringgit (MYR) ×	Jamaican Dollar (JMD) ✓	Tongan Pa'anga (TOP) ✓
Sri Lankan Rupee (LKR) ×	Jordanian Dinar (JOD) ✓	Trinidad and Tobago Dollar (TTD) ✓
Nepalese Rupee (NPR) ✓	Kenyan Shilling (KES) ×	New Taiwan Dollar (TWD) ✓
Bangladeshi Taka (BDT) ✓	Kyrgyzstani Som (KGS) ×	Ukrainian Hryvnia (UAH) ✓
Omani Rial (OMR) ✓	Cambodian Riel (KHR) ✓	Uruguayan Peso (UYU) ✓
Singapore Dollar (SGD) ✓	Comorian Franc (KMF) ×	Uzbekistani Som (UZS) ×
Afghan Afghani (AFN) ✓	Kuwaiti Dinar (KWD) ✓	Venezuelan Bolívar (VES) ×
Albanian Lek (ALL) ×	Cayman Islands Dollar (KYD) ✓	Vanuatu Vatu (VUV) ×
Armenian Dram (AMD) ✓	Kazakhstani Tenge (KZT) ✓	Samoan Tālā (WST) ×
Angolan Kwanza (AOA) ×	Lao Kip (LAK) ✓	Central African CFA Franc (XAF) ×
Argentine Peso (ARS) ×	Lebanese Pound (LBP) ✓	East Caribbean Dollar (XCD) ✓
Azerbaijani Manat (AZN) ✓	Liberian Dollar (LRD) ✓	CFP Franc (XPF) ×
Bosnia and Herzegovina Conv. Mark (BAM) ×	Lesotho Loti (LSL) ×	Zambian Kwacha (ZMW) ×
Barbadian Dollar (BBD) ×	Libyan Dinar (LYD) ✓	Zimbabwean Gold (ZWG) ×
Bulgarian Lev (BGN) ✓		

Table 35: Tick mark represents whether the symbol of that currency was used or not.

J Ethics and Risk Statement

While this work aims to expose and mitigate systemic biases in financial AI, we identify ethical considerations and potential risks associated with the dissemination of our findings.

Risk of Misinterpretation A primary concern is that the *bias scores* and *latent personas* documented in this paper (e.g., the association of certain currencies with "Malice" or "Poverty") could be misinterpreted as objective indicators of a nation's economic integrity or risk profile. We emphasize that these findings represent *representational issues* within AI models and contexts used. They do not reflect the actual financial stability, lawfulness, or economic value of the mentioned nations or their citizens.

Risk of misuse By documenting the *Syntax Gap*-specific technical instances where models fail to recognize currency symbols or acronyms-this research inadvertently identifies blind spots in current automated financial monitoring and Anti-Money Laundering (AML) systems. However, we believe that the transparency provided by this study is a prerequisite for developing a more robust and secure financial NLP infrastructure.