Profiling Cryptocurrency Influencers with Few-shot Learning

Overview for PAN Lab at CLEF 2023

Mara Chinea-Rios[1,*,†], Ian Borrego-Obrador[1,*,†], Marc Franco-Salvador[1], Francisco Rangel[1] and Paolo Rosso[2]

[1]Symanto Research, Spain
[2]Universitat Politècnica de València, Spain

Abstract
This overview presents the Author Profiling shared task at PAN2023. This year the main aim is identifying cryptocurrency influencers in social media from a low-resource perspective. We focused on English Twitter posts for three different sub-tasks: (1) *SubTask1. Low-resource influencer profiling*; (2) *SubTask2. Low-resource influencer interest profiling*; and (3) *SubTask3. Low-resource influencer intent profiling*. For this purpose, three English datasets have been provided to the participants. A total of 27 teams participated in the shared task, and we compared their performance with four baseline approaches. This overview paper reviews the approaches of the participants and presents a detailed discussion of the evaluation results.

Keywords
Author profiling, cryptocurrency influences, deep learning, low-resource, few-shot

1. Introduction

Cryptocurrencies have massively increased their popularity in recent years [1]. This complex and dynamic ecosystem encompasses a wide range of elements, including various cryptocurrencies like Bitcoin and Ethereum, smart contracts, and decentralized applications (DApps), among others.

At the current pace at which this field evolves, keeping track of its continuous development becomes a complex task. Due to the vast amount of information available nowadays, there are some who take up the challenge of riding this wave and attempting to reach its crest. The promise of independence from central authorities, increased transaction speed, cost savings (compared to the traditional financial system), and accessibility are some of the factors that attract many to this space. However, the rise of numerous frauds [2], complex language, and

CLEF 2023: Conference and Labs of the Evaluation Forum, September 18–21, 2023, Thessaloniki, Greece
*Corresponding author.
†These authors contributed equally.
mara.chinea@symanto.com (M. Chinea-Rios); ian.borrego@symanto.com (I. Borrego-Obrador); marc.franco@symanto.com (M. Franco-Salvador); francisco.rangel@symanto.com (F. Rangel); prosso@upv.edu.es (P. Rosso)
0000-0002-2313-9633 (M. Chinea-Rios); 0000-0001-7946-6601 (M. Franco-Salvador); 0000-0002-6583-3682 (F. Rangel); 0000-0002-8922-1242 (P. Rosso)

extremely high volatility also raises many concerns. In this mixture of complexity, information overload and fear is where the new influencer gold rush [3], who can shape public opinion and behavior, take the cryptocurrency realm into social media.

The study "The Impact of Social Media Influencers on Purchase Intention and the Mediation Effect of Customer Attitude" by Lim et al. [4] investigates the effectiveness of social media influencers, focusing on various factors. They found that consumer attitude mediates the relationship between these factors and purchase intention, which can be applied to the purchase and investment in cryptocurrencies. Thereby, due to the many walls that this ecosystem has and the power of influencers, we believe that many users trust social media influencers to bridge the gap in their lack of knowledge to later take investment decisions. As a consequence, profiling those influential actors becomes relevant.

Author profiling is the problem of distinguishing between classes of authors by studying how language is shared by people. This helps in identifying authors' individual characteristics, such as demographics, hate speech spreaders, among others.

With the use of the Author Profiling task in Natural Language Processing (NLP) within this field we make an approach to understand the size of an influencer, topics mentioned and interests. Their reach and ability to shape public opinion could lead to noticeable effects on the market dynamics of cryptocurrencies such as:

- *Market Movements*[1]: Influencers can cause significant fluctuations in the price and trading volume of cryptocurrencies. Positive or negative statements about a particular cryptocurrency can lead to buying or selling sprees among their followers, causing sharp rises or falls in the cryptocurrency's price.
- *Promotion and Awareness*[2]: Influencers can bring attention to lesser-known cryptocurrencies, Initial Coin Offerings (ICOs), or blockchain projects. Their endorsements can attract new investors or users, contributing to the growth and success of these projects.
- *Education*[3]: Influencers often play a role in educating the public about cryptocurrencies and blockchain technology. They can help demystify these complex topics and promote wider understanding and adoption.
- *Trust and Credibility*[3]: Influencers who are seen as experts in the cryptocurrency field can lend credibility to certain projects or cryptocurrencies. Their endorsements can build trust among potential investors.
- *Market Manipulation Concerns*[4]: There are concerns about potential market manipulation, where influencers could artificially inflate the price of a cryptocurrency for personal gain. This underscores the need for regulation and transparency in the cryptocurrency market.

It is important to note that the impact of influencers can be double-edged. While they can contribute to the growth and dynamism of the cryptocurrency ecosystem, they can also lead to increased volatility and potential risks for investors.

In this shared task, we aim to profile cryptocurrency influencers in social media from a low-resource perspective, that is, using little data. Specifically, we focus on English tweets

[1]https://tinyurl.com/investopedia-market-movements
[2]https://tinyurl.com/influencer-promotion
[3]https://tinyurl.com/investopedia-education
[4]https://tinyurl.com/manipulation-news

for three different sub-tasks: (1) *SubTask1. Low-resource influencer profiling*: profile authors according to their degree of influence (non-influencer, nano, micro, macro, mega); (2) *SubTask2. Low-resource influencer interest profiling*: profile authors according to their main interests or areas of influence (technical information, price update, trading matters, gaming, other); and (3) *SubTask3. Low-resource influencer intent profiling*: profile authors according to the intent of their messages (subjective opinion, financial information, advertising, announcement).

The remainder of this paper is organized as follows. Section 2 covers the state of the art on cryptocurrency and author profiling. Section 3 describes the created datasets and the evaluation measures, and Section 4 presents the participant's approaches. In Section 5, discuss the results achieved by the participants for each subtask and error analysis. Finally, Section 6 draws some conclusions.

2. Related Work

The purpose of this section is to provide the theoretical background. We outline the relevant works in Cryptocurrencies and author profiling.

2.1. Author Profiling

Author profiling is the problem of distinguishing between classes of authors by studying how language is shared by people. This helps in identifying authors' individual characteristics, such as age, gender, or language variety, among others. During the years 2013-2022, we addressed several of these aspects in the shared tasks organised at PAN.[5]

In 2013 the aim was to identify gender and age in social media texts for English and Spanish [6]. In 2014 we addressed age identification from a continuous perspective (without gaps between age classes) in the context of several genres, such as blogs, Twitter, and reviews (in Trip Advisor), both in English and Spanish [7]. In 2015, apart from age and gender identification, we addressed also personality recognition on Twitter in English, Spanish, Dutch, and Italian [8]. In 2016, we addressed the problem of cross-genre gender and age identification (training on Twitter data and testing on blogs and social media data) in English, Spanish, and Dutch [9]. In 2017, we addressed gender and language variety identification in Twitter in English, Spanish, Portuguese, and Arabic [10].

In 2018, we investigated gender identification on Twitter from a multimodal perspective, considering also the images linked within tweets; the dataset was composed of English, Spanish, and Arabic tweets [11].

In 2019 our focus was on profiling and discriminating bots from humans on the basis of textual data only [12] and targeting both English and Spanish tweets.

In 2020, we focused on profiling fake news spreaders [13], in two languages, English and Spanish. The ease of publishing content on social media has also increased the amount of disinformation that is published and shared. The goal of this shared task was to profile those authors who have shared some fake news in the past.

[5] To generate the datasets, we have followed a methodology that complies with the EU General Data Protection Regulation [5].

In 2021 the focus was on profiling hate speech spreaders in social media [14]. The goal was to identify Twitter users who can be considered haters, depending on the number of tweets with hateful content that they had spread. The task was set in English and Spanish.

Finally, in 2022, we focused on profiling irony and stereotype spreaders on English tweets [15]. The shared task goal was to profile highly ironic authors and those that employ irony to convey stereotypical messages, e.g. towards women or the LGTB community.

2.2. Cryptocurrencies with Machine Learning/Deep Learning

Cryptocurrencies, a digital or virtual form of currency, have become a significant area of interest in various scientific domains. The multidisciplinary nature of cryptocurrency research reflects its wide-ranging implications and the diverse challenges and opportunities it presents. Many of these domains are related with Finance, law regulations, cybersecurity etc... The domain in which we are more interested in is Computer Science (CS) and Information Technology (IT).

Within the domain we see that most works are around systems security [2], blockchain technology applications [16], predictions using time series [17, 18, 19] and scarce investigation related to social media.

Within the field we find studies around the following: The manipulation of cryptocurrencies through social media platforms, particularly "pump and dump" scams. The authors of [20] propose a computational approach to automatically identify these scams as they happen by combining information across social media platforms.

Prediction of cryptocurrency price bubbles based on social media data. The authors of [21] argue that financial price bubbles have similarities with the spread of an epidemic, as both involve a rapid increase followed by a decline. The model takes into account social media indicators such as the number of posts on a subreddit and the number of new subscribers.

Relationship between cryptocurrency price changes and topic discussion on social media. The authors of [22] use a dynamic topic modelling approach to analyze social media communication and a Hawkes model to find interactions between topics and cryptocurrency prices. They suggest that this knowledge could be built into a real-time system, providing trading or alerting signals.

Prediction of the volatile price movement of cryptocurrency by analyzing the sentiment in social media and finding the correlation between them [23].

3. Evaluation Framework

This section introduces the technical background. We outline the creation of the datasets, introduce the performance measures, baselines, and describe the software submissions procedure.

3.1. Corpus

As in previous years, a new dataset has been created from English tweets posted by users on Twitter. We built the datasets as follows: first we identified those who are crypto influencers, and next we classified their interest and intent.

We identify crypto influencer candidates with two conditions:

- user with tweets that contain the *ticker* cashtag[6] for different crypto projects e.g. $ETH, $BTC, $UNI.
- tweets with mentions in the name of the crypto projects e.g. *Ethereum, Bitcoin and Uniswap*. In the AppendixA can be found the total list of projects used.

Next, we extract the number of followers for those users. Finally, we use a follower scale to determine their influence grade. This scale is adjusted as much as possible to the most commonly accepted definition of influencer tiers[7]:

- *Non-influencer*: Individuals with a minimal social media following; typically ranging from 0 to 1,000 followers. Lacks the ability to sway opinions or impact decisions through their online presence.
- *Nano influencers*: Individuals with a small, dedicated social media following; typically ranging from 1,000 to 10,000 followers.
- *Micro influencers*: Individuals with a moderately sized social media following ranging from 10,000 to 100,000 followers. They often have a more focused and engaged audience.
- *Macro influencers*: Individuals with a substantial social media following, ranging from 100,000 to 1 million followers. They have a wide reach and may cover a broader range of topics or industries.
- *Mega influencers*: Individuals with an extensive social media following, more than 1 million followers. They often have a significant impact on popular culture and possess considerable influence across multiple platforms.

For the interest and intent datasets, we applied the following criteria after the influencer identification. For each influencer, human annotators classified the interest and intent for a random tweet sample. For the annotation process, we provide a list of labels and definitions for each task. Interest labels definition:

- *Technical information*: Post related to the technical aspects of a crypto project or asset, including infrastructure, consensus, mining, development, security, privacy, road map, compatibility etc...
- *Trading matters*: Post related to technical analysis, volatility, trading decision, risk management, and other trading-related topics.
- *Price update*: Post related to price, price speculation (e.g. "to the moon"), and price prediction, without mentioning trading-specific topics.
- *Gaming*: Post with mentions of a project or topic related to Gaming, e.g. Play2earn, GameFi, NFT drops, metaverse.
- *Other*: If none of other labels applies.

[6]Cashtags are like hashtags, but instead of using a hash sign, they use a dollar sign . They are used to tag tweets related to a specific cryptocurrency or stock. The cashtag is usually the same as the company or asset's ticker symbol on exchanges like the NYSE.

[7]https://tinyurl.com/different-tiers-of-influencers
https://twitter.com/latermedia/status/1385337617340829701
https://izea.com/resources/influencer-tiers/

Table 1
Datasets statistics including the per-class numbers of users.

Task	Partition	Total number of users per class
SubTask1	train	macro:32, mega:32, micro:30, nano:32, non-influencer:32
	test	macro:42, mega:45, micro:46, nano:45, non-influencer:42
SubTask2	train	technical information:64; trading matters:64; price update:64; gaming:64; other:64
	test	technical information:42; trading matters:112; price update:108; gaming:40; other:100
SubTask3	train	announcement:64; subjective opinion:64; financial information:64; advertising:64
	test	announcement:37; subjective opinion:160; financial information:43; advertising:52

Intent labels definition:

- *Announcement*: Post related to the announcement of events, progress, launch, partnership, or other official announcements.
- *Subjective opinion*: Post with some subjective opinion, response, or advice.
- *Financial information*: Post related to the update of price, whale signal, market, or other financial data without subjective comment.
- *Advertising*: Post related to introducing or promoting a project, account, user, website, event, etc...

The process was as follows: the tweets were annotated by three independent annotators; then, we used majority voting to select the class, and in the case of annotators' disagreement, the tweet was discarded. Also, we calculated the Fleiss Kappa [24] agreement between the annotators, obtained 0.65 for the interest task and 0.67 for the intent task. These results can be interpreted as the annotators having a substantial agreement between them in both tasks.

Table 1 presents the statistics of the datasets, and the number of users for each class. Due to the low resources task nature, the number of tweets shared with our participants is limited: for SubTask1 the maximum number of tweets is 10. To further study the low-resource author profiling, we limited the number of tweets per user in SubTask2 and SubTask3 to a single one.

3.2. Performance Measures

Since the datasets are very unbalanced, we employed macro F_1 as the official metric to measure the performance. We included also in our analysis accuracy and F_1 per class to show the participants' performance.

3.3. Baselines

We compared the participants' results with different baselines covering diverse concepts such as transfer [25] and few-shot learning [26, 27, 28]:

- *random*: labels are randomly selected with equal probability.
- *T5-large (bi-encoders) - ZS*: Zero shot (ZS) text classification employing a T5-large model [29] with bi-encoders [28].

- *T5-large (label tuning) - FS*: Few shot (FS) text classification employing a T5-large model [29] with a label-tuning training strategy [28]
- *Character n-grams with logistic regression* (user-char-lr): We use [1..5] character n-grams with a TF-IDF weighting calculated using all texts.
- *LDSE*: The Low Dimensionality Semantic Embedding (LDSE) method [30] represents documents on the basis of the probability distribution of occurrence of their tokens in the different classes, e.g., non-influencer vs. mega. The distribution of weights for a given document is expected to be closer to the weights of its corresponding class.

3.3.1. Software Submissions

Following the previous years[8], we request to the participants for software submissions. Participants submitted executables of their author profiling software instead of just the output of their software on a given test set. For the software submissions, we used the TIRA [31] experimentation platform which renders the handling of software submissions at scale as simple as handling run submissions. Using TIRA, participants employed the *Docker submission* option to load their Docker image[9], simplifying the process.

4. Overview of the Participating Systems

This year, 27 teams participated in the cryptocurrency influencer task and 12 of them submitted their working notes [10]. In this section, we analyze their approaches from four viewpoints: data preprocessing, data argumentation, features, and classification methods.

4.1. Preprocessing

Given the nature of Twitter data, 13 teams cleaned the textual contents to obtain plain text. To this end, most of them removed or masked specific elements such as URLs, user mentions, hashtags, emojis, emails, dates, money, or numbers [32, 33, 34, 35, 36, 37]. Some teams [36, 37] lowercased the tweets, removed stopwords, and treated character flooding. Also, the *hive* team removed words with less than a given number of characters.

4.2. Features Extraction

As in previous editions of the author profiling task at PAN, participants used different features, such as: (1) n-grams and (2) deep learning-based such as word2vec embeddings and BERT embeddings. Regarding n-grams, two teams (*terra* and *icon* [35]) used TF-IDF [38] n-grams. The authors of the *icon* [35] team have combined different types of n-grams, TF-IDF n-grams with word2vec [39] embeddings. For the second group, five different teams employed different transformers models to extract features. For instance, two teams (*Shashirekha* [32] and

[8] Also, we have allowed some participants to directly send us their prediction files as well as their software in the name of reproducibility.
[9] https://hub.docker.com/
[10] We received the description of the systems from the teams that did not submit their working notes.

sushiswap) employed the sentence T5 [29] model. In another case, the *ethereum* team tested different BERT [40] models, e.g. BERT, BERTweet [41]. Also, 2 teams (*waves* and *MRL-LLP*) used mpnet models trained on SNLI and MNLI datasets to get the text representations.

4.3. Data Augmentation

Some teams employed data augmentation techniques to increase the size of the training data and solve the overfitting and underfitting problems. The *holo* team used the *pegasus paraphrase*[11] model to generate variations of each original tweet while keeping its meaning. Two teams (*stellar* and *ethereum*) proposed different approaches using ChatGPT[12]. In the case of *stellar*, its approach leverages the generative capabilities of ChatGPT to generate labeled synthetic authors. Also, *ethereum* based its approach on generating new tweets belonging to the same class. Other authors [42, 43, 44] employed back translation augmentation methods. The aim was at translating all the training data into another language and then translating it back to the original language. With this technique, the author generates new data different from the original text while preserving the original context and meaning.

4.4. Approaches

Regarding the classification approaches, most participants based on fine-tuning transformer-based models [45]. We can group the transformer models used into three main groups: (1) *Encoder models*: models use only the encoder of a transformer model, e.g. BERT [40], mpnet [46]; (2) *Encoder-decoder models*: are encoder-decoder models with the goal to reconstruct [MASK] tokens in between texts, e.g. T5 [29]; and (3) *Decoder models*: are decoder models with the goal to complete continuations of texts, e.g. GPT [47], BLOOM [48]. For encoder models, some teams employed models based on BERT architecture. For instance, *neo* [34], *core* [36], *ethereum*, *magic*, *shiba-inu*, *iota*, *sushiswap*, *tron*, *iota* and *MRL-LLP* teams employed the BERT, BERTweet, CryptoBERT[13] TwHIN-BERT [49], and roBERTa [50] models. In the case of *alchemy-pay* [42], wax [44], *terra-classic*, *holo*, *vechain*, and *api3* teams included in their experiments other models like, ELECTRA [51], DistilBERT [52], Twitter-roBERTa [53], RoBERTuito [54], XML-RoBERTa [55] and DeBERTa [56], DeBERTaV3 [57]. For the second group of models, *nexo* [43], *shiba-inu*, *iota* teams tested the XLNet [58] and different T5 models e.g. T5-base, T5-large and FLAN-T5-XL [59]. In the group of decoder models, we can find that *harmony* team presented a prompt tuning approach with BLOOM model and GPT2.

Some participants have combined different approaches through stacking ensembles. The *stellar* team [33] proposed an ensemble approach that combines fine-tuned language models and natural language inference models, employing a soft voting mechanism. They tested their approach with some models, for example, BERT, CryptoBERT, FinBERT [60], roBERTa, DeBERTa. The *symbol* [61] team studied the integration of Bi-Encoders with Large Language Models (Flan-T5), to enhance the semantic representation of authors. In the case of the *dogecoin* [62] team, they focused on leveraging metric learning and Parameter-Efficient Fine-Tuning (PEFT) and

[11] https://huggingface.co/tuner007/pegasus_paraphrase
[12] https://chat.openai.com/
[13] https://huggingface.co/ElKulako/cryptobert

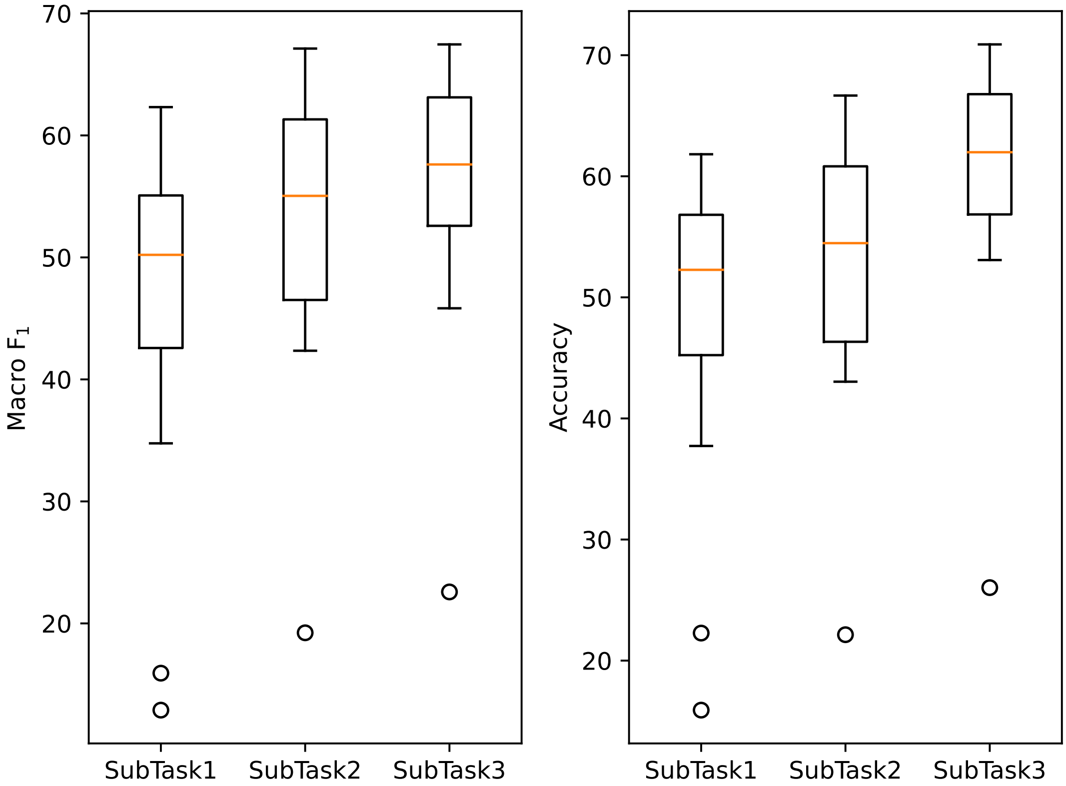

Figure 1: Distribution of results regarding Macro F_1 and Accuracy for each subtask.

an application of PEFT techniques to CryptoBERT and BLOOM-1b1 models. Participants of the *Abhinav.Kumar* [37] team proposed to combine the bodies of two pre-trained transformer models, the models employed were TwHIN-BERT and roBERTa.

With respect to other deep learning approaches, the *waves* team employed CNN with sentence embedding and *hive* team used Prototypical networks [63] and MiniLM-L12.

Only a few participants approached the task with traditional approaches. *Shashirekha* [32] team used Support Vector Machine (SVM), *icon* [35] team used Random Forest, Logistic Regression and SVM. Additionally, 5 teams (*shiba-inu*, *vechain*, *tron*, *terra*, *ethereum*) complemented the analysis with the following approaches: Gradient Boosting, Stochastic Gradient Descent (SGD) classifier, KNeighbors classifier, Gradient Boosting Classifier, Gaussian NB, Multi-layer Perceptron classifier, etc...

5. Evaluation and Discussion of the Results

In this section, we present the results of the shared task, as well as analyze the most common errors made by the teams.

Table 2
Ranking for the Subtask1: Low-resource influencer profiling in terms of macro F_1 performance.

Pos	Team	Macro F_1
1	holo	**62.32**
2	terra-classic	61.14
3	stellar [33]	58.44
4	MRL-LLP	57.44
5	magic [64]	57.14
6	vechain	55.51
7	neo [34]	55.10
8	waves	55.06
9	iota	54.43
10	hive	52.94
11	symbol [61]	52.31
12	dogecoin [62]	50.80
13	shiba-inu	50.38
14	Abhinav.Kumar [37]	50.21
	LDSE	50.20
15	tron	50.13
	t5-large (label tuning) - FS	49.34
16	api3	49.18
17	terra	48.74
18	harmony	47.93
19	ethereum	46.68
20	sushiswap	46.64
21	alchemy-pay [42]	38.51
22	nexo [43]	38.34
23	Shashirekha [32]	37.92
24	wax [44]	37.62
	user-char-lr	35.25
25	core [36]	34.76
26	solana	15.92
	random	15.92
27	icon [35]	12.90
	t5-large (bi-encoders) - ZS	12.76

5.1. Tasks Ranking

Tables 2 and 3 present the overall performance of the participants for each task. The results are shown in terms of Macro F_1. The best result for Subtask1 was obtained by *holo* team (62.32) following the approach of fine-tuning the RoBERTuito [54] model. For Subtask2 the best result was obtained by the *stellar* [33] team (67.12) with data augmentation techniques and a method of transformer-based models ensemble. In this case, we observed how the *stellar* team with the data augmentation strategy outperforms the second position in three points. In Subtask3, the *terra-classic* team obtained the best results with fine-tuning DeBERTaV3 [65]. Comparing the participants' results with our baseline, we can show that around 46% of the final submissions

Table 3
Ranking for the Subtask2 and Subtask3

(a) Ranking for the Subtask2: Low-resource influencer interest profiling in terms of macro F_1 performance.

Pos	Team	Macro F_1
1	stellar [33]	**67.12**
2	iota	64.55
3	terra-classic	63.15
4	MRL-LLP	62.00
5	neo [34]	61.63
6	symbol [61]	61.21
7	vechain	60.16
8	shiba-inu	58.47
9	holo	57.50
	t5-large (label tuning) - FS	56.48
10	magic [64]	55.68
11	harmony	54.41
	user-char-lr	52.95
12	dogecoin [62]	51.72
13	hive	51.48
14	tron	49.77
15	Shashirekha [32]	46.66
16	api3	46.07
	LDSE	44.92
17	terra	44.60
18	core	43.47
19	waves	42.35
	t5-large (bi-encoders) - ZS	33.34
	random	20.81
20	sushiswap	19.23

(b) Ranking for the Subtask3: Low-resource influencer intent profiling in terms of macro F_1 performance.

Pos	Team	Macro F_1
1	terra-classic	**67.46**
2	shiba-inu	66.15
3	symbol [61]	65.83
4	MRL-LLP	65.74
5	stellar [33]	64.46
6	api3	63.12
7	holo	61.81
8	magic [64]	61.62
9	vechain	60.28
	user-char-lr	60.21
	t5-large (label tuning) - FS	59.91
10	hive	59.08
11	neo [34]	57.62
12	ethereum	55.94
13	core	55.34
14	terra	54.83
15	tron	53.43
16	dogecoin [62]	52.59
	LDSE	51,96
17	iota	50.62
18	Shashirekha [32]	50.42
19	waves	49.21
20	harmony	45.83
	t5-large (bi-encoders) - ZS	32.71
21	sushiswap	22.58
	random	18.41

outperformed our best baseline for each subtask. In addition, only one submission performed worse than the *random* baseline (see Table 2). The traditional approaches could not outperform the deep learning-based, positioned on average from the twelve positions of the ranking. Finally, the best systems could achieve an improvement of up to 10% absolute macro F_1 score compared to our best baselines.

As can be seen in Figure 1, the results are shown in terms of Macro F_1 and accuracy for each subtasks. In the case of SubTask1 the median results are 50.21 Macro F_1 and 52.27 accuracy with an inter-quartile range of 12.50% and 11.59%. Although the median is similar for SubTask 2 and Subtask3 (55.05 vs. 57.62) Macro F_1 with an inter-quartile range of 14.80% and 10.53%. Nevertheless, in the case of Subtask1, the standard deviation is higher (12.01 vs. 10.97-9.97), due to the higher number of outliers.

Figure 2 shows the results for each class and tasks in terms of F_1 score. In the case of SubTask1

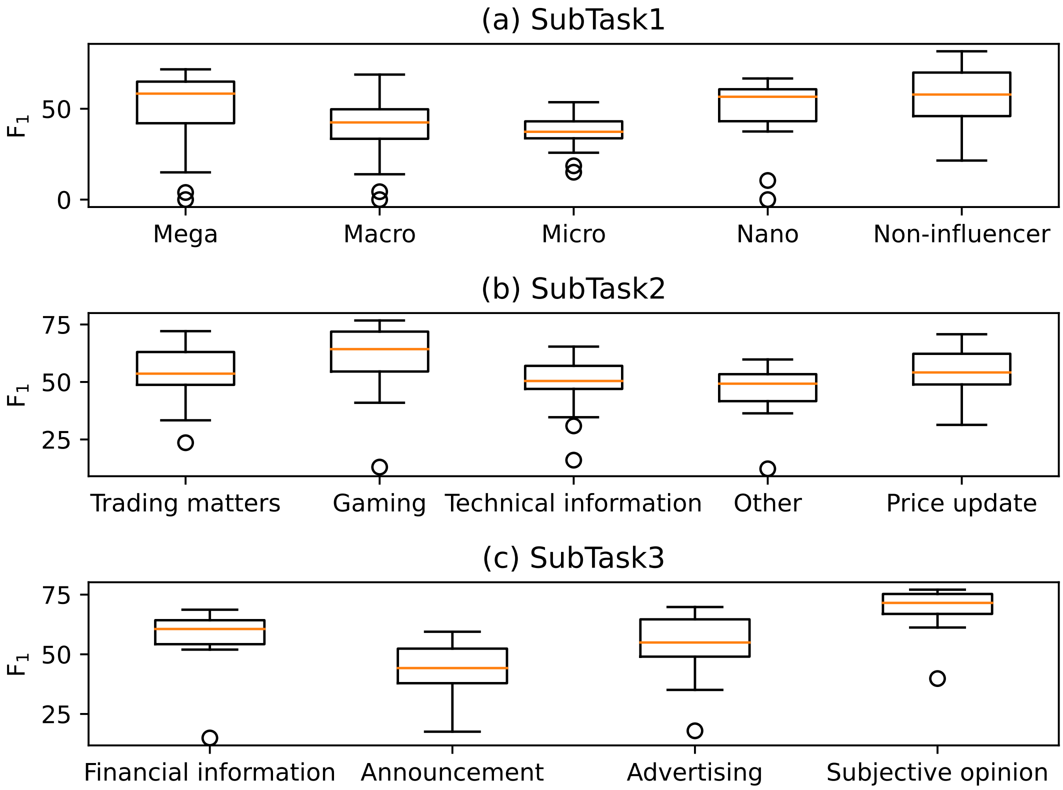

Figure 2: Distribution of results in terms of F_1 for each class.

(see Figure 2 (a)), the standard deviation values are high in almost all classes (mega: 20.52, macro: 17.43, micro: 9.06, nano:15.67 and non-influencer: 15.31), we can interpret these results that some systems had low predictions results and produced a high number of outliers increasing the sparsity. The micro class obtained the lowest median 37.33 with an inter-quartile range of 9.33% In SubTask2 (see Figure 2 (b)), the standard derivation on average is 11.79, and the median on average is 54.35. We can show similar results with SubTask3 (see Figure 2 (c)), the standard deviation on average is 10.87, and the average median of 57.83.

5.2. Error Analysis

Figures 3, 4 and 5 show the confusion matrices. We have aggregated all the participants' predictions, except baselines, and plotted the respective confusion matrices. Figure 3 shows the aggregated confusion matrix for SubTask1. We can see that macro and mega are easier to identify with respect to the other class but they get confused with each other (30% and 43%). We can see a similar behavior between nano and micro with lower values (34% and 27%) and non-influencer could be confused with nano and micro influencer (12% and 20%). The aggregated confusion matrix for SubTask2 is shown in Figure 4. The most difficult class

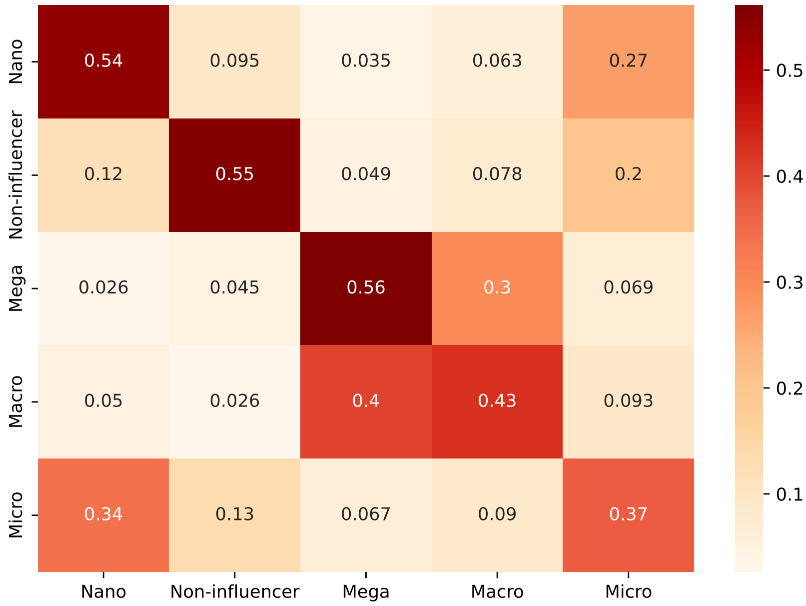

Figure 3: Aggregated confusion matrix for SubTask1.

to classify is the other class (44%) and the easier is gaming (77%). Also, the trading matters class is mainly confused with the price updates class (23%). Figure 5, we have aggregated the predictions for SubTask2. In this case, the financial information class got the highest value (79%), from which we could interpret that authors with this intent are more recognizable. On the contrary, the advertising class is confused with the announcement class (24% and 15%).

6. Conclusion

In this paper we have presented the results of the 11th International Author Profiling Shared Task at PAN 2023, hosted at CLEF 2023. This edition focused on English texts, and the participants had to profile cryptocurrency influencers in Twitter, from a low-resource perspective. Specifically, we requested to profile the influence degree, the interests, and intents of the influencers.

The participants used different features to address the task, mainly: (1) n-grams and (2) deep learning-based such as classical word embeddings (e.g. word2vec) and more modern transformer-based ones (e.g. BERT). Concerning the classification part, most of the teams approached the task with deep learning techniques, being transformer-based classifiers (BERT and DeBERTA) the most popular.

The best results for profiling the influencer's degree (SubTask1) have been obtained by the *holo* team with (62.32 Macro F_1) with the approach of fine-tuning the RoBERTuito transformer-based model. The best results for the interests subtask (Subtask2) have been obtained by the *stellar* [33] team (67.12 Macro F_1) with data augmentation technique employed GPT and a method which combined a BERT model with an NLI model. Finally, the best results for the intents subtask

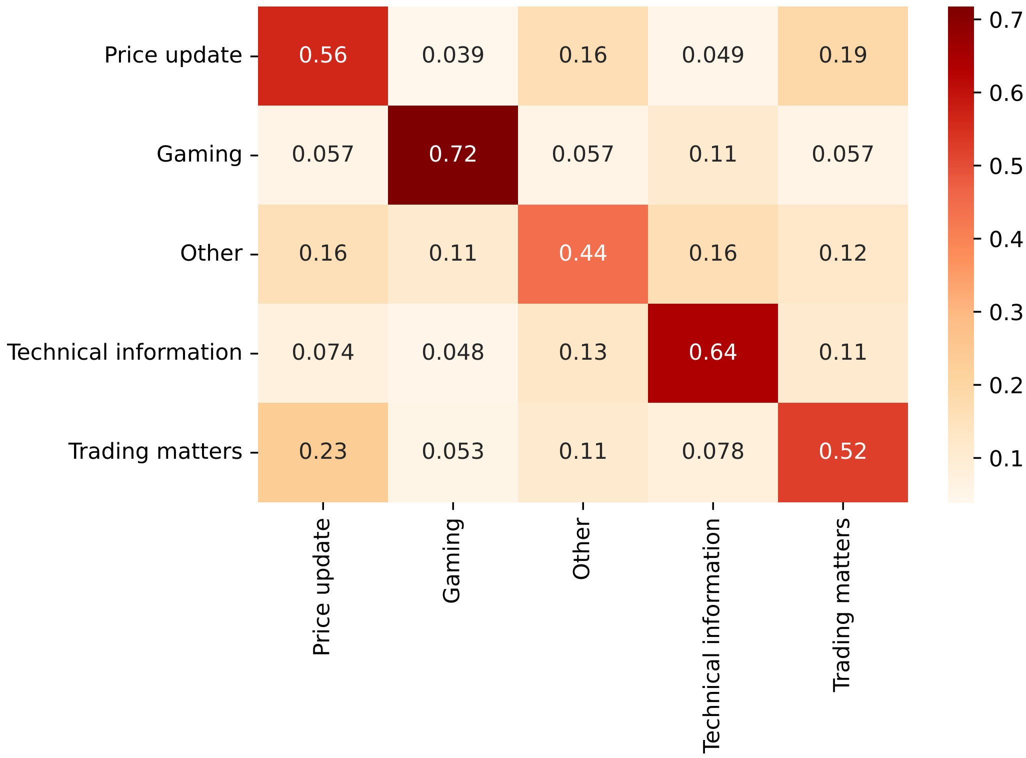

Figure 4: Aggregated confusion matrix for SubTask2.

(Subtask3) have been obtained by the *terra-classic* team (67.46 Macro F_1) following the approach of fine-tuning a transformer-based model, in this case DeBERTaV3.

The error analysis per SubTask1 shows that the highest confusion is from macro to mega (40%). Similar behavior is shown between micro and nano (37%). The non-influencer category could be confused with micro influencer (20%). In the case of SubTask2 the error analysis shows that the most difficult class to classify by the participant systems is the other class (44%) and the easier to identify is gaming class (77%). For SubTask3, the authors' most recognizable intent comes from the financial information class (79%), and advertising and announcement can be confused (24% and 15%).

Looking at the results and the error analysis, we can conclude that: (1) it is possible to automatically identify cryptocurrency influencers and their interest and intent with acceptable Macro F_1 (while there is room for improvement); (2) transformer-based models outperform traditional approaches for the task of cryptocurrency influencers. This highlights their potential for low-resource author profiling scenarios.

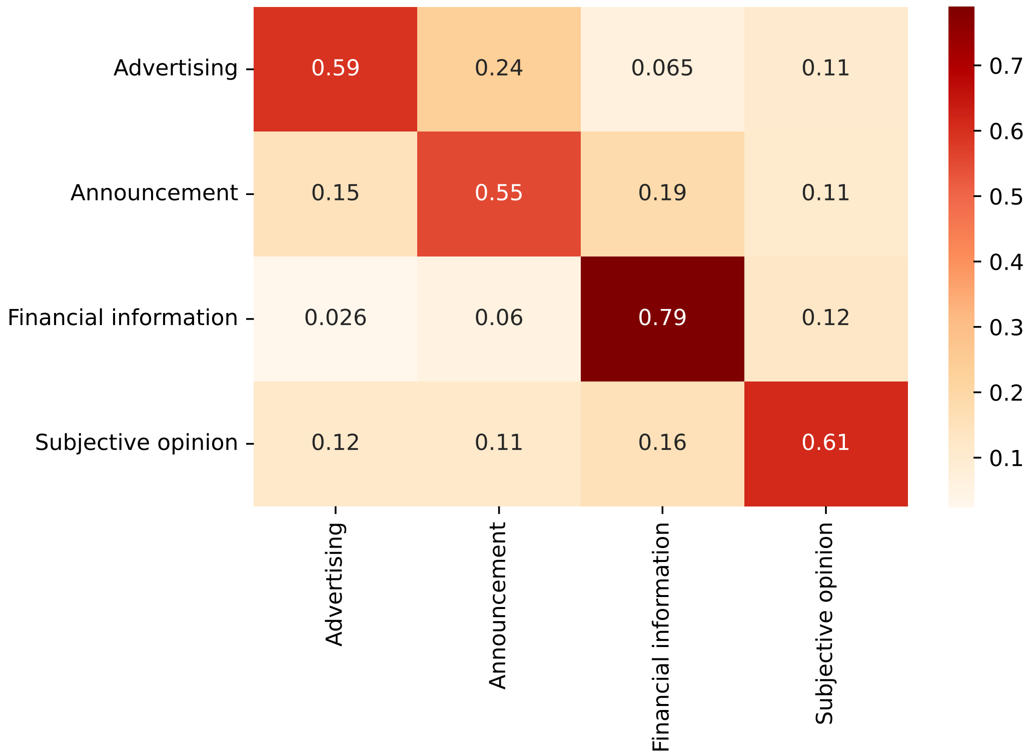

Figure 5: Aggregated confusion matrix for SubTask3.

Acknowledgments

First, our special thanks go to the 27 participants teams for providing high-quality submissions for this shared task. It is encouraging to see the high acceptance of this new shared task focused on low-resource cryptocurrency influencer author profiling.

The work from Symanto has been partially funded by the Pro^{2}Haters - Proactive Profiling of Hate Speech Spreaders (CDTi IDI-20210776), the XAI-DisInfodemics: eXplainable AI for disinformation and conspiracy detection during infodemics (MICIN PLEC2021-007681), the OBULEX - *OBservatorio del Uso de Lenguage sEXista en la red* (IVACE IMINOD/2022/106), and the ANDHI - ANomalous Diffusion of Harmful Information (CPP2021-008994) R&D grants. The work of Paolo Rosso was in the framework of the FairTransNLP research project (PID2021-124361OB-C31), funded by MCIN/AEI/10.13039/501100011033 and by ERDF, EU A way of making Europe.

References

[1] R. Sawhney, S. Agarwal, V. Mittal, P. Rosso, V. Nanda, S. Chava, Cryptocurrency bubble detection: A new stock market dataset, financial task & hyperbolic models, in: Proceedings

of the Conference of the North American Chapter of the Association for Computational Linguistics, 2022, pp. 5531–5545.
[2] H. Chen, M. Pendleton, L. Njilla, S. Xu, A survey on ethereum systems security: Vulnerabilities, attacks, and defenses, ACM Computing Surveys (CSUR) 53 (2020) 1–43.
[3] S. Khamis, L. Ang, R. Welling, Self-branding,'micro-celebrity'and the rise of social media influencers, Celebrity Studies 8 (2017) 191–208.
[4] X. J. Lim, A. Radzol, J. Cheah, M. W. Wong, The impact of social media influencers on purchase intention and the mediation effect of customer attitude, Asian Journal of Business Research 7 (2017) 19–36.
[5] F. Rangel, P. Rosso, On the implications of the general data protection regulation on the organisation of evaluation tasks, Language and Law / Linguagem e Direito 5 (2019) 95–117.
[6] F. Rangel, P. Rosso, M. Moshe Koppel, E. Stamatatos, G. Inches, Overview of the author profiling task at PAN 2013, in: CLEF 2013 Labs and Workshops, Notebook Papers, 2013.
[7] F. Rangel, P. Rosso, I. Chugur, M. Potthast, M. Trenkmann, B. Stein, B. Verhoeven, W. Daelemans, Overview of the Author Profiling Task at PAN 2014, in: CLEF 2014 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2014.
[8] F. Rangel, P. Rosso, M. Potthast, B. Stein, W. Daelemans, Overview of the Author Profiling Task at PAN 2015, in: CLEF 2015 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2015.
[9] F. Rangel, P. Rosso, B. Verhoeven, W. Daelemans, M. Potthast, B. Stein, Overview of the Author Profiling task at PAN 2016: Cross-genre evaluations, in: CLEF 2016 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2016.
[10] F. Rangel, P. Rosso, M. Potthast, B. Stein, Overview of the 5th author profiling task at PAN 2017: Gender and language variety identification in Twitter, 2017.
[11] F. Rangel, P. Rosso, M. Montes-y-Gómez, M. Potthast, B. Stein, Overview of the Author Profiling Task at PAN 2018: Multimodal Gender Identification in Twitter, in: CLEF 2019 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2018.
[12] F. Rangel, P. Rosso, Overview of the author profiling task at pan 2019: Bots and gender profiling., in: CLEF 2019 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2019.
[13] F. Rangel, A. Giachanou, B. Ghanem, P. Rosso, Overview of the Author Profiling Task at PAN 2019: Profiling Fake News Spreaders on Twitter, in: CLEF 2020 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2020.
[14] F. Rangel, G. L. De-La-Peña-Sarracén, B. Chulvi, E. Fersini, P. Rosso, Profiling Hate Speech Spreaders on Twitter Task at PAN 2021, in: CLEF 2021 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2021.
[15] O.-B. Reynier, C. Berta, R. Francisco, R. Paolo, F. Elisabetta, Profiling irony and stereotype spreaders on twitter (irostereo) at pan 2022, in: CLEF 2021 Labs and Workshops, Notebook Papers. CEUR Workshop Proceedings, 2022.
[16] T. Dursun, F. Birinci, B. Alptekin, I. Sertkaya, O. Hasekioglu, B. Tunaboylu, S. Zaim, Blockchain technology for supply chain management, in: Proceedings of the Industrial Engineering in the Internet-of-Things World, 2022, pp. 203–217.
[17] I. E. Livieris, E. Pintelas, S. Stavroyiannis, P. Pintelas, Ensemble deep learning models for forecasting cryptocurrency time-series, Algorithms 13 (2020) 121.

[18] M. Mudassir, S. Bennbaia, D. Unal, M. Hammoudeh, Time-series forecasting of bitcoin prices using high-dimensional features: a machine learning approach, Neural Computing and Applications (2020) 1–15.

[19] S. A. Apostu, M. Panait, L. Vasa, C. Mihaescu, Z. Dobrowolski, Nfts and cryptocurrencies—the metamorphosis of the economy under the sign of blockchain: A time series approach, Mathematics 10 (2022) 3218.

[20] M. Mirtaheri, S. Abu-El-Haija, F. Morstatter, G. Ver Steeg, A. Galstyan, Identifying and analyzing cryptocurrency manipulations in social media, IEEE Transactions on Computational Social Systems 8 (2021) 607–617.

[21] R. C. Phillips, D. Gorse, Predicting cryptocurrency price bubbles using social media data and epidemic modelling, in: Proceedings of the Symposium Series on Computational Intelligence, 2017, pp. 1–7.

[22] R. C. Phillips, D. Gorse, Mutual-excitation of cryptocurrency market returns and social media topics, in: Proceedings of the International Conference on Frontiers of Educational Technologies, 2018, pp. 80–86.

[23] X. Huang, W. Zhang, X. Tang, M. Zhang, J. Surbiryala, V. Iosifidis, Z. Liu, J. Zhang, Lstm based sentiment analysis for cryptocurrency prediction, in: Proceedings of the International Conference Database Systems for Advanced Applications, 2021, pp. 617–621.

[24] J. L. Fleiss, Measuring nominal scale agreement among many raters., Psychological bulletin 76 (1971) 378.

[25] K. Weiss, T. M. Khoshgoftaar, D. Wang, A survey of transfer learning, Journal of Big data 3 (2016) 1–40.

[26] M. Chinea-Rios, T. Müller, G. L. De la Peña Sarracén, F. Rangel, M. Franco-Salvador, Zero and few-shot learning for author profiling, in: Proceedings of the International Conference on Applications of Natural Language to Information Systems, 2022, pp. 333–344.

[27] T. Müller, G. Pérez-Torró, A. Basile, M. Franco-Salvador, Active few-shot learning with fasl, in: Proceedings of the International Conference on Applications of Natural Language to Information Systems, 2022, pp. 98–110.

[28] T. Mueller, G. Pérez-Torró, M. Franco-Salvador, Few-shot learning with siamese networks and label tuning, in: Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2022, pp. 8532–8545.

[29] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, Exploring the limits of transfer learning with a unified text-to-text transformer, The Journal of Machine Learning Research 21 (2020) 5485–5551.

[30] F. Rangel, P. Rosso, M. Franco-Salvador, A low dimensionality representation for language variety identification, in: Proceedings of the Computational Linguistics and Intelligent Text Processing, 2016, pp. 156–169.

[31] M. Potthast, T. Gollub, M. Wiegmann, B. Stein, Tira integrated research architecture, Information Retrieval Evaluation in a Changing World: Lessons Learned from 20 Years of CLEF 41 (2019) 123–160.

[32] G. Kavya, H. Asha, B. Fazlourrahman, S. H. Lakshmaiah, Profiling cryptocurrency influencers with sentence transformers, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.

[33] E. Villa-Cueva, J. M. Valles-Silva, A. P. Lopez-Monroy, F. Sanchez-Vega, J. R. Lopez-Santillan, Integrating fine-tuned language models and entailment-based approaches for low-resource tweet classification, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[34] Z. Li, Z. Han, J. Cai, Z. Huang, S. Huang, L. Kong, Clem4pci: Profiling cryptocurrency influencers with few-shot learning via contrastive learning and ensemble model, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[35] H. Muslihuddeen, P. Sathvika, S. Sankar, S. Ostwal, A. Dr. Kumar M, Profiling cryptocurrency influencers using few-shot learning, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[36] D. Y. Espinosa, G. Sidorov, Using bert to profiling cryptocurrency influencers, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[37] A. Kumar, A. A. Saeed, L. H. M. Trinh, Profiling cryptocurrency influencers with few shotlearnings, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[38] G. Salton, C. Buckley, Term-weighting approaches in automatic text retrieval, Information processing & management 24 (1988) 513–523.
[39] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, arXiv e-prints (2013).
[40] J. D. M.-W. C. Kenton, L. K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of North American Chapter of the Association for Computational Linguistics, 2019, pp. 4171–4186.
[41] D. Q. Nguyen, T. Vu, A. Tuan Nguyen, BERTweet: A pre-trained language model for English tweets, in: Proceedings of the Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 2020, pp. 9–14.
[42] M. Siino, M. Tesconi, I. Tinnirello, Profiling cryptocurrency influencers with few-shot learning using data augmentation and electra, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[43] M. Siino, I. Tinnirello, Xlnet on augmented dataset to profile cryptocurrency influencers, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[44] F. Lomonaco, M. Siino, M. Tesconi, Text enrichment with japanese language to profile cryptocurrency influencers, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.
[45] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Advances in neural information processing systems 30 (2017) 1–10.
[46] K. Song, X. Tan, T. Qin, J. Lu, T.-Y. Liu, Mpnet: Masked and permuted pre-training for language understanding, Advances in Neural Information Processing Systems 33 (2020)

16857–16867.
[47] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, in: Advances in Neural Information Processing Systems, volume 33, 2020, pp. 1877–1901.
[48] T. Le Scao, A. Fan, C. Akiki, E. Pavlick, S. Ilić, D. Hesslow, R. Castagné, A. Sasha Luccioni, F. Yvon, M. Gallé, et al., Bloom: A 176b-parameter open-access multilingual language model, arXiv e-prints (2022) arXiv–2211.
[49] X. Zhang, Y. Malkov, O. Florez, S. Park, B. McWilliams, J. Han, A. El-Kishky, Twhin-bert: A socially-enriched pre-trained language model for multilingual tweet representations, arXiv e-prints (2022) arXiv–2209.
[50] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, arXiv e-prints (2019).
[51] K. Clark, M.-T. Luong, Q. V. Le, C. D. Manning, Electra: Pre-training text encoders as discriminators rather than generators, in: Proceedings of the International Conference on Learning Representations, 2020, pp. 109–120.
[52] V. Sanh, L. Debut, J. Chaumond, T. Wolf, Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, arXiv e-prints (2019).
[53] F. Barbieri, J. Camacho-Collados, L. Neves, L. Espinosa-Anke, Tweeteval: Unified benchmark and comparative evaluation for tweet classification, arXiv e-prints (2020).
[54] J. M. Pérez, D. A. Furman, L. A. Alemany, F. M. Luque, Robertuito: a pre-trained language model for social media text in spanish, in: Proceedings of the Language Resources and Evaluation Conference, 2022, pp. 7235–7243.
[55] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, É. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, in: Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2020, pp. 8440–8451.
[56] P. He, X. Liu, J. Gao, W. Chen, Deberta: Decoding-enhanced bert with disentangled attention, in: Proceedings of the International Conference on Learning Representations, 2020, pp. 1050–1061.
[57] P. He, J. Gao, W. Chen, DeBERTav3: Improving deBERTa using ELECTRA-style pre-training with gradient-disentangled embedding sharing, in: Proceedings of the International Conference on Learning Representations, 2023, pp. 140–151.
[58] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, Q. V. Le, Xlnet: generalized autoregressive pretraining for language understanding, in: Proceedings of the International Conference on Neural Information Processing Systems, 2019, pp. 5753–5763.
[59] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, E. Li, X. Wang, M. Dehghani, S. Brahma, et al., Scaling instruction-finetuned language models, arXiv e-prints (2022).
[60] D. Araci, Finbert: Financial sentiment analysis with pre-trained language models, arXiv e-prints (2019).
[61] H. B. Giglou, M. Rahgouy, J. D'Souza, T.-H. B. A. Oskuee, Milad Molazadeh, C. D. Seals,

Leveraging large language models with multiple loss learners for few-shot author profiling., in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.

[62] R. Labadie Tamayo, A. M. Sarvazyan, Reshape or update? metric learning and fine-tuning for low-resource influencer profiling, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.

[63] J. Snell, K. Swersky, R. Zemel, Prototypical networks for few-shot learning, Advances in Neural Information Processing Systems 30 (2017) 4077–4087.

[64] I. Ferri-Molla, J. Santamaria-Jorda, Profiling cryptocurrency influencers with few-shot learning, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, CEUR-WS.org, 2023.

[65] P. He, J. Gao, W. Chen, Debertav3: improving deberta using electra-style pre-training with gradient-disentangled embedding sharing, in: Proceedings of the International Conference on Learning Representations, 2023, pp. 1–23.

A. Crypto Projects

For our data crawling, we monitored 122 crypto projects. Table 4 lists in alphabetical order the project name and the corresponding tickers.

Table 4
List of crypto projects employed.

Name	ticker	Name	ticker	Name	ticker
AAVE	$AAVE	Filecoin	$FIL	Polygon	$MATIC
Algorand	$ALGO	Flow	$FLOW	Quant Network	$QUAN
Alice	$ALICE	FRAX	$FRAX	Quickswap	$QUICK
Amp	$AMP	Frax Share	$FXS	Rally	$RLY
Anchor	$ANCT	FTX Token	$FTT	Raydium	$RAY
ApeCoin	$APE	Gala	$GALA	Rocket Pool	$RPL
Aptos	$APT	Gnosis	$GNO	Shiba INU	$SHIB
Arweave	$AR	Handshake	$HNS	Solana	$SOL
Audius	$AUDIO	Harmony	$ONE	Spell Token	$SELL
Avalanche	$AVAX	Hedera	$HBAR	Stacks	$STX
Axie Infinity	$AXS	Helium	$HNT	Star Atlas	$POLIS
Balancer	$BAL	Huobi Token	$HT	Stellar	$XLM
Bancor	$BNT	Illuvium	$ILV	STEPN	$GMT
Basic Attention Token	$BAT	Immutable X	$IMX	Storj	$STORJ
Binance USD	$BUSD	Internet Computer	$ICP	SushiSwap	$SUSHI
Bitcoin Cash	$BCN	Keep Network	$KEEP	Synapse	$SYN
Braintrust	$BTRST	KuCoin	$KCS	Synthetix	$SNX
Cardano	$ADA	Kusama	$KSM	Terra	$TERA
Celo	$CELO	Lido DAO	$LDO	Tether	$USDT
Chainlink	$LINK	Litecoin	$LTC	Tezos	$XTZ
Chia	$XCH	Livepeer	$LPT	The Graph	$GRT
Chiliz	$CHZ	LooksRare	$LOOKS	The Sandbox	$SAND
Compound	$COMP	Loopring	$LRC	Theta Network	$THETA
Convex Finance	$CVX	MAGIC	$MAGIC	ThorChain	$RUNE
Cosmos	$ATOM	Maker	$MKR	Tron	$TRX
Cream Finance	$CREAM	MINA	$MINA	Truebit	$TRU
Cronos	$CRO	Mirror	$NUL	TrueFi	$TRU
Curve	$CURVE	Monero	$XMR	TrueUSD	$TUSD
DAI	$DAI	MultiversX	$EGLD	Trust Wallet Token	$TWT
Decentraland	$MANA	NEAR Protocol	$NEAR	UMA	$UMA
DODO	$DODO	NEO	$NEO	Uniswap	$UNI
Dogecoin	$DOGE	NEXO	$NEXO	Unus Sed Leo	$LEO
dYdX	$DYDX	NFT Worlds	$WRLD	USD Coin	$USDC
Enjin Coin	$ENJ	Oasis Network	$ROSE	VeChain	$VET
EOS	$EOS	Olympus DAO	$OHMINU	VENUS	$XVS
Ethereum	$ETHE	Optimism	$OP	Vulcan Forged PYR	$PYR
Ethereum Classic	$ETC	Orchid	$OXT	Waves	$waves
Ethereum Name Service	$ENS	PancakeSwap	$CAKE	Yearn Finance	$YFI
Fantom	$FTM	ParaSwap	$PSP	Yield Guild Games	$YGG
FaraLand	$FARA	Perpetual Protocol	$PERP	Zcash	$ZEC
Fetch AI	$FET	Polkadot	$DOT		