Possibilities of Automatic Detection of Reactions to Frustration in Social Networks

Yulia Kuznetsova[a], Natalia Chudova[a], Vladimir Salimovsky[b], Daria Sharypina[c] and Dmitry Devyatkin[a]

[a] *Federal Research Center "Computer Science and Control" RAS, 44-2 Vavilova str., Moscow, 119333, Russia*
[b] *Perm State University, 15 Bukireva str., Perm, 614068, Russia*
[c] *Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, 141701, Russia*

Abstract

This study aims to create a method for the reliable detection of frustration-derived reactions in social media texts. Based on the results obtained earlier while automating the categorization of Rosenzweig Picture-Frustration Study responses, a method was created to automatically classify the reactions to frustration found in social network posts and comments. The experiment results show that the E, E', M reactions can be reliably detected with fair precision and recall, although we have obtained lower F_1 scores for other reactions because those classes are very small. The results prove that Rosenzweig's types of frustrating responses can also be applied to the study of social media behavior. Moreover, the language used to express a particular reaction to frustration is not related to the content of the situation. The elaborated method currently works only for two genres: answers in Rosenzweig's test and comments or posts in social media. Recognizing the types of reactions to frustration in other genres may require a new algorithm adjustment.

Keywords

Reactions to frustration, Rosenzweig Picture-Frustration Study, machine learning, social networks

1. Introduction

Rosenzweig Picture-Frustration Study (RPFS) was created in 1945 by S. Rosenzweig and long entered the "golden fund" of psychodiagnostics. Decades of using the technique in many countries, including Russia, have shown its high effectiveness in identifying personal peculiarities of responding to obstacles and accusations. The method is considered semi-projective but not challenging to master, so, for example, most psychology students already master it in their psycho-diagnostics practical classes. This is probably because the test has good clarity of assessing the ways of responding to an obstacle proposed by the author, and these reactions' language expressions are distinct. Examples of typical responses are often given in the guidelines for using the technique, and some authors provide detailed lists of such examples (for Russian-speaking subjects, see, for example, [1]). However, there is still no systematic description of speech reactions to frustration in psychology.

In the last decade, the elaboration of automatic text analysis tools and text classification methods based on machine learning has become very intensive. With these methods, one could adapt well-established psycho-diagnostic techniques for use in the information society, where people produce a large flow of texts. Social networks have become data banks of hundreds of millions of users, including their speech reactions to various negative and positive circumstances.

In work [2], we solved the problem of elaborate a tool for automatically classifying the responses of subjects in the RPFS. A corpus of 462 RPFS protocols was collected, and the psychologist processed these texts identifying subjects' reactions to frustration. At the next stage, the linguist worked with the

IMS 2021 - International Conference "Internet and Modern Society", June 24-26, 2021, St. Petersburg, Russia
EMAIL: devyatkin@isa.ru (A.5)
ORCID: 0000-0001-9380-4478 (A.1); 0000-0002-0811-725X (A.5)

marked-up corpus, and the linguistic descriptions were formalized and applied to construct a feature description of text fragments. Finally, the feature descriptions were used to build a classifier of reactions to frustration utilizing machine learning methods. It was found that the resulting linguistic patterns form a high-level feature description of text fragments that allows for high completeness (R is not less than 0.8) identifying statements related to different types of reactions. Four of the nine types of reactions: M, M', I, E, can be reliably distinguished (F1> 0.7) without considering the context of the statements. It was noted that for psychology, the technology of linguistic patterns acts, on the one hand, as a means of professional reflection, and on the other as a tool for verifying the data of projective text techniques, and allows us to develop tools for automatic analysis of the respondents' speech, including online discussions. It was suggested that further work with the elaborated linguistic patterns could be aimed at testing the hypothesis of their universality concerning frustrating situations. In other words, the results allowed us to assume that such speech responses will occur in any frustrating situation, and not just in the ones presented in RPFS.

A team of psychologists, linguists, and artificial intelligence specialists tested the effectiveness of the elaborated algorithm for automatically detecting reactions to frustration in the texts of online discussions. Note that the technology of the linguistic pattern developed by the authors [3] involves the participation of a psychologist (or any socio-humanitarian scientist: historian, sociologist, political, etc., interested in the text researching), a linguist, and an artificial intelligence specialist. The pattern technology can be considered a new method for modeling the reasoning of an expert who evaluates texts within the framework of the categorical scheme adopted in their discipline.

The second step of our study is devoted to testing the algorithm's effectiveness in social media texts, which would make it possible to automatically classify subjects' responses in the Rosenzweig test to a particular type. In that step we deal with the following research questions.

1. Can the Rosenzweig's types of frustrating responses also be found in social media behavior? The speech design of such reactions as an accusation, complacency, justification, or the willingness to solve the problem independently is not too diverse. Therefore, the linguistic means used by users of social networks should be about the same as the means used by subjects when performing the Rosenzweig test.
2. Is the linguistic means used to express a particular reaction to frustration are not related to the content of the frustrating situation? If so, regardless of the subject of discussion, communicants who describe their frustrating reactions use universal mechanisms of expression (speech patterns), making it possible to identify the type of reaction to frustration in a wide range of contexts.

2. Related work

Attempts to analyze the emotional states of social networks users are popular, including active efforts to identify the features of texts written in a state of stress, frustration, or grief [4-6]. In our study, we try to identify text features of social network user frustration, comparing posts by calm and well-being users and messages by the same users in a state of tension [7]. However, we have not found any studies devoted to detecting reactions to frustration identified by S. Rosenzweig. Plenty of works are devoted to detecting sentiment, mood, or affect in social networks, which seems quite close in terms of valuable features and approaches. Those studies often consider only shallow lexical features; for example, Thelwall presents [4] TensiStrength, a system to detect the strength of stress and relaxation expressed in social media text messages. TensiStrength uses a lexical approach with lists of terms related to stress and relaxation. Those terms are synonyms for stress, anxiety, and frustration and terms related to anger and negative emotions because stress can be a response to negative events and can cause negative emotions. Thelwall claims that the effectiveness of TensiStrength depends on the nature of the messages, with the texts that are rich in stress-related terms being particularly problematic. The experiment results show that TensiStrength works well enough to be applied for applications that need to use stress information.

The paper [8] presents a study more complex case. In this paper, the researchers propose a method to detect sarcasm. Sarcasm is a form of text in which individuals state the opposite of what is implied. The researchers utilize theories from behavioral and psychological studies to construct a behavioral modeling framework tuned for detecting sarcasm. That presumes using more complex features.

Namely, they observe that sarcastic texts sometimes have a specific structure wherein the author's views are expressed in the message's first few words. Simultaneously, in the later parts, a description of a particular scenario is put forth. To reveal possible syntactic patterns arising from such text construction, researchers use the POS tags of the first three words and the last three words in the texts. They also include the position of the first sentiment-loaded word and the first affect-loaded word as a feature. To capture differences in syntactic structures, they consider POS tags present in the message. Namely, they build a probability distribution over the POS tags present in the current text and POS tags in past messages and use the Jensen-Shannon divergence value between the two distributions as a feature. They also use lexical density, which is the fraction of information-carrying words present in the text (nouns, verbs, adjectives, and adverbs).

Complex linguistic features are also used in the paper [9] for hate-speech detection. Researchers collected more than 2M texts, comparing discussion actors around neutral topics to those more likely to be hate-related. They combine word embeddings, sentiment, and emotional features, lexis, and POS tags and apply bidirectional Long-Short-Term Memory (LSTM) [10] because the training corpus was big enough.

There are also several works related to cognitive distortion detection. The primary problem here is the lack of training corpora. For example, the paper [11] presents an approach to classify text into one of 15 distortion categories. They compared several machine learning-based classifiers, such as Logistic regression, SVM, recurrent neural networks (Gated Recurrent Units) [12], gradient boosting on decision trees (XGBoost) [13]. The best-performing model is again logistic regression because the dataset was relatively small.

An example of the practical application of sentiment, mood, or affect detection methods is presented in the thesis [14]. In this thesis, Primetshofer uses sentiment analysis to detect frustrating conversation situations. He claims that it is helpful for chat-bot systems. Such a system should check the sentiment of the user's input message and clarify the current situation. The proposed method analyses several types of features like lexemes, POS tags, and syntax dependences. Then it uses a machine learning technique for analyzing the opinion. He uses the method to detect the moments when systems stumble and fail to answer the request. They require a human agent's help and intelligence; in this situation, a transition from the machine to a human agent is one of the core features.

To summarize, frustration-derived reaction detection requires representative datasets to train the classification models with rich contextual linguistic features. However, the creation of such a dataset involves a lot of manual data collection and annotation. Unfortunately, the most complex classification models lack interpretability, which is important for psychodiagnostics. Therefore, in this work, we focus on context-aware but pretty simple models and classification features, which can be easily interpreted and verified.

3. Pikabu Frustration Dataset

We selected the text material among the messages posted on the entertainment site Pikabu.ru without considering the subject of the discussion [15]. Namely, the discussions included in the analyzed corpus were selected according to the maximum representation of the types of authors' reactions to frustration. The experimental dataset contains 528097 sentences manually annotated with 11 classes. Two categories have been added to the nine original types of frustration response: 1) informing (in their comments, people sometimes, along with one of the reaction types, give quite detailed information about the discussed situation, attracting their knowledge in the field of law, history, technology, etc.); and 2) instruction (information about how such a conflict situation should, or can, or must be resolved in a particular society or community).

In total, the texts of 1943 unique post and comment authors were analyzed. After marking up the building, psychologists identified 3,490 cases of responding to frustration, including: E: 1579, e: 200, E': 390, I: 64, i: 129, I': 79, M: 147, m: 41, M': 201, informing: 528, instructing: 132.

Those messages are related to various controversial topics. We collected those messages in such a way to make each class (reaction) multi-topical to avoid the use of topic-related lexis by the frustration-derived reactions classifier. We have been guided by the linguistic description from section 4 when labeling. Because of the texts' nature, the dataset is severely imbalanced. Namely, the 'E' class is more than a dozen times bigger than the second-largest class. In addition to pure texts, the dataset contains information about relationships between the messages (post-comment), making it possible to catch the context for each message.

4. Linguistic patterns

Below we present the linguistic descriptions, which have been created during the "manual" analysis of the RPFS subject responses and then served as guidelines for the automatic analysis of online discussion texts. In total, the linguist compiled 60 rules, some of which were a set of more specific rules, and some a cliché list or a word list.

Language indicators of E'-reactions:

1. Impersonal sentences with the main member expressed by a predicative. The predicative set of the group E’ is outlined quite clearly: *грустно, досадно, жалко, обидно, паршиво, печально, плохо* (*sad, annoying, miserable, upset, lousy, bad*), etc. (the linguist provided the AI specialist with a complete list of words in this role were found in 426 protocols of RPFS).
2. Nominative sentences belonging to an evaluative-beingness group. When the main member expressed by a noun in the nominative case, emotional-expressive particles is often used.
3. Interjections used as separate independent utterances. Their set in the class E’ is finite and is provided by the linguist in the form of a short word list.
4. Sentences with the subject expressed by the personal pronoun *я* (*I*) (less often by a noun in the nominative case), and the predicate expressed by: a) a short adjective, b) a short passive participle. The adjective or participle has the semantics of a negative emotional state. The attribution of these sentences to the class E’ or to the class E is determined by the totality of the individual subject's reactions.
5. Two-part sentences with the subject expressed by the anaphoric pronoun *это* (*this*), and the predicate expressed by a negative-evaluative a) noun, b) adverb (the adverbs are the same as in impersonal sentences). Differentiation of such statements between E’ and E can be carried out according to the composition of predicates.
6. Two-part and one-part sentences with the predicate expressed by the conjugated form of a verb denoting a negative emotional state or if it has the semantics of a negative evaluation.

Language indicators of E-reactions:

1. Definitely-personal sentences with the main member (predicate) expressed by: a) a full-significant verb in the imperative mood (often with the negative particle *не* (*not*), b) the connective verb *быть* (*to be*) in the imperative mood *будь, будьте* (-), and the nominal part by an adjective in the short form or in the form of the comparative degree.
2. Sentences with the compound verb predicate including: a) the modal verb *мочь* (*can*) in the subjunctive mood мог бы, могли бы (*could*) and an adjacent infinitive, b) impersonal-predicative words with the modal meaning: *нужно, надо, можно, следует, стоит, нельзя, хватит* (*need, have to, may, should, worth, cannot, enough*), and an adjacent infinitive.
3. Sentences with the predicate represented by words of the negative-evaluative semantics (nouns, adjectives, participles, adverbs) or the negative-action semantics (verb sentences). These include:

a) N1-N1: The subject is a noun or a personal pronoun in the nominative case, the predicate is a noun in the nominative case. In the role of the subject, the anaphoric pronoun *это* (*this*) is often used, too;

b) N1-Adj: The subject is a noun in the nominative case, the predicate is a full or short form of an adjective;

c) Inf-Adv: The subject is an infinitive, the predicate is an adverb. The subject is regularly represented by the anaphoric pronoun *это* (*this*);

d) N1-Part1short: The subject is a noun or a personal pronoun in the nominative case, the predicate is a short passive participle with the semantics of a negative emotional state as a result of an action;

e) N1: Nominative sentence. The main member expressed by a noun in the nominative case;

f) N1-Vf: The subject is a noun or a personal pronoun in the nominative case, the predicate is a conjugated form of a verb.

There is a group of sentences with verbs of a negative emotional state: *злит, бесит, расстроило* (*makes/drives smb. angry, crazy, sad*).

4. Sentences of different structures with the contextual meaning of confrontational negation. Their indicator is the negative particle не (no). They are used to: a) refute an opinion or position of the interlocutor; b) express a reproach or claim.

5. Rhetorical questions or rhetorical exclamations with the pronominal adverbs *как, почему (how, why*), and its synonyms *на каком основании, по какой причине* (*on what purpose, for what reason*); *куда, зачем (where, why*), and its synonym *к чему* (*for what*); *причем, где, сколько* (*what, where, how long*); the pronouns *кто, что, какой* (*who, what, which*), etc. expressing the semantics of denial, indignation, perplexity. Before the pronominal adverb, the conjunctions *а, и* (*but, and*), and the particle *да* (*let*) are often used.

6. Interrogative sentences of the confrontational semantics with the particle *что ли* (*or what*).

7. Communicative fragments, i.e., "ready-to-use pieces of language material" (B.M. Gasparov). In principle, they can be specified in a list.

Language indicators of e-reactions:

1. Interrogative sentences with the compound verbal predicate including the modal verb *мочь* (*can*) in the subjunctive mood *мог бы, не мог бы, могли бы, не могли бы* (*could, could not, would, would not*) and an adjacent infinitive. In these sentences, the subject is expressed by the second person pronoun *ты, вы* (*you*). In some cases, it can be omitted. This is an indirect speech act expressing a request in the form of a question. Meanwhile, the narrative sentences of this structure (graphically not having the question mark at the end) no longer express the request, but a condemnation, a reproach and enter the class E (see above the E-pattern).

2. Interrogative sentences with the compound verb predicate including the modal verb *(с)мочь* (*can*) in the indicative mood *можете, можешь* (*can*), and an adjacent infinitive. The subject (in some cases it is omitted) is again usually expressed by the second person pronoun *ты, вы* (*you*). This indirect speech act, as in the previous case, expresses a request.

3. Interrogative sentences with the main member expressed by impersonal predicative words *можно* in the meaning "*possible*", *нельзя* (*impossible*), and an adjacent infinitive. Sometimes the predicative unit *я могу* (*I can*) acts as a synonym for a word. The speech act expresses: a) a request to perform the action indicated by the infinitive; b) a request to allow or to permit the speaker himself to perform the action indicated by the infinitive (it is assumed that the addressee will give one's permission, i.e., on one's part will perform the required action); c) a request to provide the information the speaker needs.

4. Interrogative sentences with the semantics of motivation expressed by the future tense form of a verb in the presence of the introductory words *может быть, может, возможно* (*can be, maybe perhaps*), or less often the word *можно* in the meaning "*allowed, permitted*".

5. Sentences with the verb predicate in the imperative mood, expressed by a combination of the particle пусть (let) with a verb of the 3rd person singular and plural future tense.

6. Definitely-personal sentences with the predicate expressing a motivation for joint action by combining the particles *давай, давайте* (*give, let's*) with a verb of the 1st person plural future tense.

7. Verbs *скажи, скажите, подскажи, подскажите* (*tell me*) as key words in the main part of a complex sentence with the explanatory subordinate part.

8. Sentences with the predicate expressed by a verb of speech (most often the verb in the performative use, i.e., in use meaning the performance of an action called the verb), and an adjacent infinitive as the complement.

9. Definitely-personal sentences in which the main member (predicate) is expressed by a full-significant verb in the imperative mood. Whether such sentences enter the class e or the class E, is determined by the totality of the subject's reactions.

10. Sentences with the compound verb predicate, expressed by the predicative *должен* (*must*), and an adjacent infinitive. Whether such sentences enter the class e or the class E, is determined by the totality of the subject's reactions.

11. Sentences with the main member expressed impersonally-predicative words *надо, нужно, необходимо, следует, придется, стоит, пора, лучше* (*need, have to, ought to, necessary, should, be to, it's time for, better*), and an adjacent infinitive.

12. Reproducible phrases like *будьте добры* (*please*) + infinitive*, мне нужно* (*I need*), etc. (the complete list has been created).

Language indicators of I'-reactions:

1. Subjective-predicative sentences with the predicate expressed by a verb with the negative particle не (no), if there is the combination *все равно* (*all the same*) as a particle.
2. Compound sentences with a subordinate explanatory, in which the main part is an impersonal sentence with the predicates *радует, отрадно, хорошо, отлично* (*please, gratifying, good, great*).
3. Other cases of using the word *хорошо* (*good*) as a predicate.
4. Sentences with the predicate *уберечь* (*save*).
5. Sentences with the predicate *рад* (*glad*).
6. Using the comparative degree of the adverb *хороший* (*good*) – i.e., *лучше* (*better*).
7. Using the combination of conditional conjunctions *если, раз* and the particle *уж (if so, once so*).
8. Using the negative conjunction *зато* (*but, although*).

Language indicators of I-reactions:

1. Predicative units directly expressing regret, guilt *извини(те), прости(те), прошу прощения, прошу меня извинить, извиняюсь, приношу свои извинения, виноват, сожалею, мне нет прощения* (*I'm sorry, excuse me, I apologize, beg your pardon, forgive me, I regret, I am beyond redemption*). Such units account for 55% of all recorded reactions.
2. Predicative units that are reproduced in a ready-made form and semantically diverse reporting the on the unintentional nature of the committed action, or its recklessness, or the readiness to correct what happened, to be punished, or the intention to correct yourself, or contain a promise not to commit such actions again (a list of 13 cliches).
3. The idea of the unintentional nature of the committed action is expressed by the verbal predicates *хотеть, знать, заметить, (по)думать, (у)видеть (want, know, notice, think, see*) in the form of the past tense with the negative particle *не* (*no*), as well as the reproducible *phrases [вышло] по неосторожности* (*it came out by negligence/accident*).
4. Sentences with the subject expressed by the pronoun *я* (*I*; can be omitted), and the predicate by a short or full adjective. The predicate contains lexemes denoting the subject's features, the manifestation of which he indirectly apologizes.
5. Impersonal sentences with the predicatives *жаль, жалко, стыдно, неловко* (*sorry, pity, ashamed, embarrassing*).
6. Commissives (speech acts by which the speaker assumes certain obligations) containing the performative *обещаю (I promise*).
7. Statements and single words expressing the speaker's agreement with the charge against him *вы правы, да, действительно, согласен, признаю* (*you are right, yes, indeed, I agree, I admit*).

Language indicators of i-reactions:

1. Subjective-predicative sentences N1–Vf, in which the function of the subject is performed by a 1st person singular or plural pronoun of the *я, мы* (*I, we*), and the function of the predicate is performed by a verb in the form of the future tense. Sentences of this type form the absolute majority of speech i-reactions.
2. Subjective-predicative sentences with the subject expressed by a 1st person pronoun, and the compound verb predicate expressed by the modal verb *мочь* (*can*) (less often by the verb *хотеть* (*want*)) in the form of the 1st person singular or plural of the present tense *могу, можем, хочу* (*I can, we can, I want*) and an adjacent infinitive.
3. Subjective-predicative sentences with the subject expressed by a 1st person pronoun, and the compound verb predicate, expressed by the short adjective *должен* (*must*) and an adjacent infinitive.
4. Impersonal sentences with the main member expressed by the predicative *надо, стоит, придется* (*necessary, should, have to*) and an adjacent infinitive.
5. The particles *да, ладно, хорошо* (*yes, okay, well*) expressing an agreement with the interlocutor, or an intention to give in to him.
6. The particle *(ну) что ж* in the meaning "*I have to agree*".

Language indicators of M'-reactions:

Indicators are communicative fragments (speech units reproduced in the ready-to-use form) *ничего, ничего страшного, не страшно, не беда, всё в порядке, всё обошлось, всё нормально, всё хорошо,* *я в порядке, всё отлично, не беспокойтесь, не переживайте, без проблем, бывает, ладно* (*nothing, nothing terrible, not scary, it doesn't matter, everything is in order, everything worked out, everything is fine, everything is well, I'm fine, don't worry, no problem, it happens, okay*) etc. (the full list contains 44 cliches). In this case, it is impractical to highlight syntactic models, since we are dealing mainly with cliched speech reactions. Often a phrase contains two different communicative fragments: *Все нормально, ничего страшного* (*It's okay, don't worry*). It should also be taken into account that in M'-reaction, the frequency of the use of speech etiquette formulas (words and phrases *спасибо, спасибо за беспокойство*, *до свидания, всего доброго* (*thank you, thank you for your concern, goodbye, all the best*), etc.) is increased.

Language indicators of M-reactions:

1. Using the predicates *случаться* and *бывать* in the meaning "*to happen*", usually in the impersonal use, sometimes as the predicate of a two-part sentence.
2. Sentences with negation containing words with the root *-вин-: вина, виноват(ый), винить* (*a guilt, guilty, to blame*). In most cases, these are subjective-predicative sentences with the subject expressed by a personal or negative pronoun.
3. Communicative fragments (reproducible fragments of language matter) expressing the ideas that a) everything happened by accident, without intent or because of circumstances have arisen; b) the reason for everything is fate, predestination; c) nothing can be changed.
4. Sentences with a subject expressed by the pronoun *я* (*I*), and a predicate expressed by the verb *понимать* (*understand*) implicitly conveys the idea that the speaker has no complaints about the interlocutor.
5. Subjective-predicative nominal sentences N1-Adj, i.e., with the subject expressed by a noun in the nominative case, and the predicate expressed by an adjective or a pronoun-adjective. In context, they implicitly express an idea that the cause of a trouble is in circumstances, and not in the person actions *Жизнь такая; Это часы такие* (*This is the life; It's a useless watch*).
6. Subjective-predicative nominal sentences N1–N1 with lexically matching subject and predicate *Дети есть дети; Правила есть правила* (*Children are children; Rules are rules*). In the context, they imply the absence of claims against anyone.
7. Definitely-personal sentences with the verb-predicate *не волнуйтесь, не огорчайтесь* (*don't worry, take it easy*).
8. Statements with the word *ничего* (*nothing*) as a particle in the expression of consent, acceptance of what happened, as well as with the phrase *ничего страшного* (*nothing terrible*).
9. Using the words *(ну и) ладно, что ж* (*All right, well*), the phraseology is *Бог с ним* (Literally, *God with it = Well, never mind, whatever*) when expressing consent or concession.

Language indicators of m-reactions:

1. Sentences with the verb-predicate *подождать* (*wait*) in the form of the 1st person plural future tense подождем (*we'll wait*), sometimes in combination with the particle *давай(те) подождем (let's wait*). Such sentences regularly include dimensives (components with the meaning of measure). They account for 25% of all obtained m-reactions, and this is quite natural.
2. Statements with introductory words expressing the uncertainty *возможно, видимо, вероятно, может, может быть, должно быть* (*perhaps, apparently, probably, maybe, should be, possible*), or less often the confidence *наверняка* (*for sure*) in what is being said, and the motive (and therefore the implicit meaning) of these statements is to explain and justify someone's actions. Since the speaker is looking for an explanation for the actions of a third person, the noun *причина* (*reason*) is regularly used. Thus, the marker of these statements is one of the specified introductory words in the presence of the noun reason.
3. Using temporatives with the meaning "after a while": *позже, позднее, в следующий раз, в другой раз, при встрече* (*later, sometime later, next time, another time, when see*).

5. Methods of the frustration detection

Because of the severe class imbalance, we had to use down-sampling. For this, we took the two most voluminous classes ('E' and 'inf'), and for each of them, we built random subsets of objects, which were smaller than the original class size (30K samples for the class E, and 20K samples for the class "inf"). We also applied the following pre-processing for the texts. First, all texts are divided into tokens-words, tokens are reduced to lower case, punctuation is removed.

Second, we extracted the pattern-based features, high-level features built with the linguistic patterns from section 4. The cornerstone of the patterns is the relational-situational model of a clause, a heterogeneous semantic network (HSN) of syntaxemes with a specific structure [16]. We define the context-free patterns as a list of HSNs that matches parts of the clauses with implementations of a particular reaction. Those HSNs can be partially defined if some linguistic feature (lemma, grammar case, etc.) of a syntaxeme is not essential for the classification.

We have built 60 such patterns based on the cognitive-communicative action markers revealed for the specific reactions by psychologists and linguists. The pattern-based feature-set generation is a process, which analyzes the message clauses with the dependency and SRL parsers [17] and matches them with the context-free patterns. If a pattern contains lemmas, we utilize the pre-trained Fasttext model [18, 19] to perform the fuzzy comparison between the text and the pattern; therefore, misspellings and synonyms can be processed. Eventually, each clause can be represented with a binary vector, which encodes if the clause matches the patterns.

Further, we build lexical features. For each text, we remove stop words and build *tf-idf* vectors of token unigrams and bigrams. Eventually, models are trained on the obtained vectors. We trained pretty simple models with a sliding window approach to catch the context of the messages. Those models are based on linear support vector machines (SVM), logistic regression, Random Forest, and Gradient boosting to classify the reaction types.

6. Experiment Results

All models were trained with weights to balance the classes. We used F_1-micro metric and stratified 5-fold cross-validation to select the values of the hyperparameters, including the size of the window for context extraction. The best result is 0.73 F_1-micro score for the Gradient boosting trained on the combination of the patterns and lexical features. The pure lexis provides 0.48 F_1-micro only.

Models were evaluated on the holdout sub-corpus with the same class distribution as the original dataset. We use standard classification scores to assess detection reliability, which are precision, recall, and F1-score. Let describe them in more detail.

- *tp* is the number of correctly detected sentences containing particular type of reactions;
- *fp* is the number of sentences that do not contain particular type of reactions but are incorrectly assigned to this type by the classifier;
- *fn* is the number of the sentences with particular type of reaction incorrectly assigned to other types.

The precision (P) is the share of correctly identified sentences from all sentences marked by the classifier as a particular type of reaction. Recall (R) - the proportion of correctly identified sentences of a particular type of reaction and F_1-score – harmonic mean of precision and recall (1).

$$P = \frac{tp}{tp + fp}, R = \frac{tp}{tp + fn}, F_1 = \frac{2PR}{P + R} \tag{1}$$

Table 1 shows the results for the classification on the test subset for the best model and features combination (for well represented classes).

Table 1
Performance of the reaction detection on the labeled dataset

Reaction	P	R	F_1-score
E	**0.77**	**0.89**	**0.89**
E'	**0.67**	0.45	0.54
M	**0.74**	**0.69**	**0.71**
M'	**0.62**	**0.72**	**0.67**
e	0.48	0.28	0.35
i	**0.64**	0.36	0.46

It is worth noting that the method extracts *E, M, M’* types pretty accurately. The precision scores are fair for all the types except *i*. Fig. 1 presents the final distribution of the predictions.

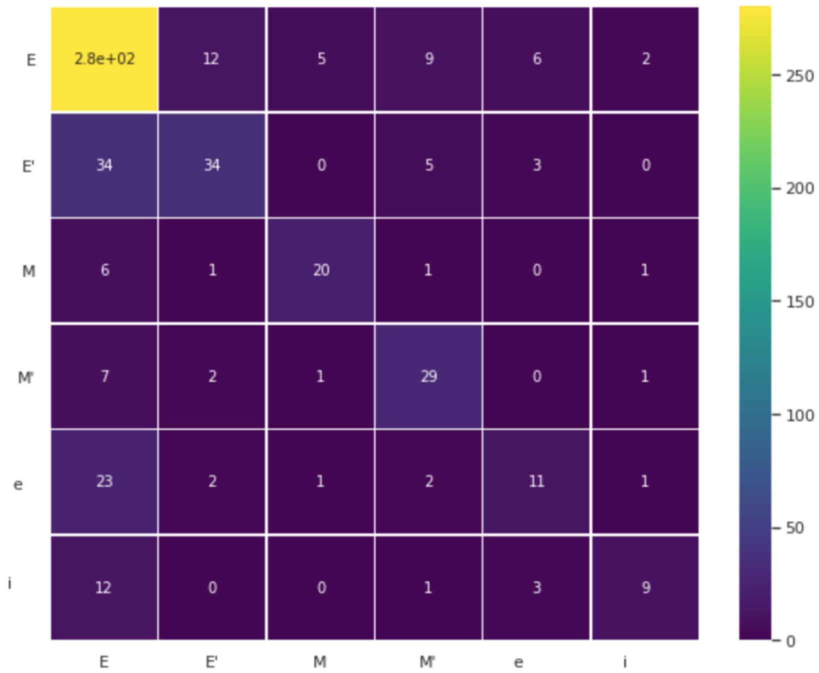

Figure 1: Confusion matrix

The most share of misclassification is related to the E/E’ and E/e pairs. We believe that is partly because of imbalance of the data; therefore the accuracy for those types can be improved when we extend the corpus. The obtained result could be achieved on such a complex material as social media text due to two conditions: the accuracy of Saul Rosenzweig's typology and the linguistic "formulas" as a "tutor" in machine learning.

7. Conclusion

Based on the results obtained earlier while automating the categorization of Rosenzweig Picture-Frustration Study responses, the algorithm was created to automatically classify the reactions to frustration found in social network posts and comments. The methods of machine learning applied to the corpus of network discussions previously marked up by psychologists allowed us to obtain a tool for automatically distinguishing reactions to frustration in posts and comments of social network users. However, we should point to the constraints of the genre of the analyzed text as the limitations of the created method. The method currently works only for two genres: answers in Rosenzweig's test and comments or posts in social media. Recognizing the types of reactions to frustration in online counseling texts or nonfiction or fictional texts may require a new algorithm adjustment.

It is worth noting that the corpus of texts on which the training took place is not large enough and is collected from discussions posted only in one of the popular Russian-language social networks. Further development of the created tool will require expanding the corpus by attracting material from other social networks.

8. Acknowledgements

This study is supported by Russian Foundation for Basic Research. Grant No 18-29-22047 mk.

9. References

[1] L. A. Yasyukova, Frustration test P.Rosenzweig, Methodological guide, IMATON, 2004.

[2] D. A. Devyatkin, S. N. Enikolopov, V. A. Salimovsky, N. V. Chudova, Speech reactions to frustration: automatic categorization. Psikhologicheskie Issledovaniya [=Psychological Studies], 2021. To appear.

[3] D. A. Devyatkin, V. A. Salimovsky, N. V. Chudova, Cognitive approach to computer analysis of scientific texts. In: Proceedings of the Eighth International Conference on Cognitive Science: Abstracts of reports. Svetlogorsk, October 18-21, A. K. Krylov, V. D. Solovyov (Eds), Institute of Psychology of the Russian Academy of Sciences, 2018, pp. 1064-1067.

[4] M. Thelwall, TensiStrength: Stress and relaxation magnitude detection for social media texts. J. Information Processing & Management, 53(1), 2017, pp. 106-121.

[5] J. Brubaker, F. Kivran-Swaine, L. Taber, H. Gillian, Grief-Stricken in a Crowd: The Language of Bereavement and Distress in Social Media. In: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media. The AAAI Press, Palo Alto, California, 2012, 42-49.

[6] D. Carr, D. Umberson: The Social Psychology of Stress, Health, and Coping. In: Handbook of Social Psychology, 2013, pp. 465-487. https://doi.org/10.1007/978-94-007-6772-0_16.

[7] S.N. Enikolopov, A.K. Kovalev, Y.M. Kuznetsova, E.V. Starostina, N.V. Chudova, Feature characteristic of texts written in a state of frustration, Bulletin of Moscow University, Series 14. Psychology, 3, 2019. pp. 66-85. (in Russ.)

[8] A. Rajadesingan, R. Zafarani, and H. Liu, Sarcasm detection on twitter: A behavioral modeling approach. In: Proceedings of the eighth ACM international conference on web search and data mining, February, 2015, pp. 97-106.

[9] D. Chatzakou, I. Leontiadis, J. Blackburn, E. De Cristofaro, G. Stringhini, A. Vakali, N. Kourtellis, Detecting cyberbullying and cyberaggression in social media, ACM Transactions on the Web (TWEB), 13(3), 2019, pp. 1-51.

[10] S. Hochreiter, J. Schmidhuber: Long short-term memory, Neural computation, 8(9), 1997, pp. 1735-1780.

[11] B. Shickel, S. Siegel, M. Heesacker, S. Benton, P. Rashidi, Automatic Detection and Classification of Cognitive Distortions in Mental Health Text. In: arXiv preprint arXiv:1909.07502, 2019.

[12] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling. In: arXiv preprint arXiv: 1412.3555, 2014.

[13] J. H. Friedman, Greedy function approximation: a gradient boosting machine. J. Annals of statistics, 2001, pp. 1189-1232.

[14] M. Primetshofer, Detection and Handling of Frustrating Conversation Situations in a Text-Based Chatbot System., Master Thesis, 2019.

[15] Pikabu social network, 2021. URL: https://pikabu.ru.

[16] G. S. Osipov, I. V. Smirnov, I. A. Tikhomirov, Relational-situational method for text search and analysis and its applications, Scientific and Technical Information Processing 37(6), 2010, 432-437.

[17] D. Larionov, A. Shelmanov, E. Chistova, and I. Smirnov, Semantic role labeling with pretrained language models for known and unknown predicates, In: Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019), 2019, pp. 619-628.

[18] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, Enriching word vectors with subword information, Transactions of the Association for Computational Linguistics 5, 2017, pp. 135-146.

[19] V. Benko, V.P. Zakharov, Very large Russian corpora: new opportunities and new challenges. In: Computational linguistics and intellectual technologies, 2016, pp. 79-93.