Recognition of Genuine Polish Suicide Notes

Maciej Piasecki
Wrocław University
of Science and Technology
Wrocław, Poland
`maciej.piasecki@`
`pwr.edu.pl`

Ksenia Młynarczyk
Wrocław University
of Science and Technology
Wrocław, Poland
`ksenia.mlynarczyk@`
`gmail.com`

Jan Kocoń
Wrocław University
of Science and Technology
Wrocław, Poland
`jan.kocon@ pwr.edu.pl`

Abstract

In this article we present the result of the research on the recognition of genuine Polish suicide notes (SNs). We provide useful method to distinguish between SNs and other types of discourse, including counterfeited SNs. The method uses a wide range of word-based and semantic features and it was evaluated using Polish Corpus of Suicide Notes, which contains 1244 genuine SNs, expanded with a manually prepared set of 334 counterfeited SNs and 2200 letter-like texts from the Internet. We utilised the algorithm to create the class-related sense dictionaries to improve the result of SNs classification. The obtained results show that there are fundamental differences between genuine SNs and counterfeited SNs. The applied method of the sense dictionary construction appeared to be the best way of improving the model.

1 Introduction

Suicide is a tragedy for the victim and also for their close ones. It is also the third leading cause of violent death among people aged 15 to 44 (Holmes et al., 2007), see also (Gomez, 2014; World Health Organisation, 2014). The reasons for such an act and the mental state of a victim are not open by itself to external observation. However, very often the last language utterances are left in a form of *suicide notes* (henceforth SNs). Such recorded utterances create a unique opportunity to come closer to the way of thinking of someone at risk, and to construct a model of the specific language which is used by people in such a state of mind. The analysis can go in two possible directions: firstly recognition of suicide letters among other types of writing and secondly identification of the

features that are characteristic for suicide notes and can provide some insight on the person committing suicide. Both of them are closely correlated and for both development of classification methods to separate SNs from other types of writing is crucial.

Linguistic analysis in (Zaśko-Zielińska, 2013) showed that such differences are mainly of the semantic and pragmatic nature. Moreover, SNs have personal character and varied length but with the dominance of short notes. In addition examples of genuine SNs are available in small data sets. A distinction between genuine SNs and texts intentionally written in such a way that they are meant to resemble SNs may provide crucial evidence for finding out the intrinsic features of SNs, if there are any. Our goal was to develop classification method for the recognition of genuine SNs among other types of texts with a special focus given for sorting out text only resembling SNs, especially counterfeited SNs. The method should analyse a wide range of linguistic features and be a good basis for the automated identification of features that make SNs so specific.

As we can expect that the differences between suicide notes and other types of discourse can be mainly of the semantic and pragmatic natures, we wanted to expand the corpus analysis beyond the simple statistical analysis of word occurrence.

2 Related Works

The study of SNs has a long tradition of qualitative analysis from the point of view of linguistics and clinical psychology (Shneidman and Farberow, 1957). There have also been attempts to perform statistical analysis (Gomez, 2014), e.g. Pennebaker and Chung (2011) used the frequency of verbal elements in a narrative which express a certain mood or sentiment.

Pestian et al. (2008) pioneered automated recognition of differences between genuine SNs

and SNs written by volunteers, known as elicited. They worked with a sample of 33 genuine and 33 elicited items. Descriptive features were based on text segmentation and morpho-syntactic tagging only. The trained classifiers achieved performance above the 50% precision baseline. The data set was small and the number of shared words among notes limited, so Pestian et al. (2008) also manually annotated texts with emotion labels from a limited set of categories; that improved the result. Besides the classification itself, they were interested in features which appeared to be significant for the classification. The significant features became a starting point for a linguistic and psychological analysis of the authors of the genuine notes. Pestian et al. (2010) worked with the same 66 SNs. They computed such text characteristics as part of speech, information, readability scores and parse information, and performed manual classification: "trainees accurately classified notes 49% of the time" and "mental health professionals accurately classified notes 63% of the time". The expanded set of features gave a 78% accuracy of automatic classification, but no semantic or emotionally motivated features were considered. The words selected as features seem to be specific only to this particular set of documents. The words arising from feature selection seem to be quite accidental and specific to this particular set of documents, but not specific to SNs.

Matykiewicz et al. (2009) extended that work to a much larger collection of more than 600 genuine SNs. Words frequent enough in SNs were put into overlapping classes with respect to the emotions contained in the Linguistic Inquiry and Word Count tool (LIWC) (Pennebaker et al., 2001). It is worth noting that emotion labels were assigned to words (lemmas), not to word senses (lexical meanings). Matykiewicz et al. (2009) concentrated on document clustering; elicited SNs were not considered. Authors tried to find features which distinguish genuine SNs from other forms of personal communication. As the background, they used posts to different newsgroups, selecting those newsgroups which seemed to be thematically close to the suicide discourse: talk.politics.guns, talk.politics.mideast, talk.politics.misc and talk.religion.misc. There were good clustering results (above 90% of cluster purity), but the background corpus did not include elicited SNs, which seem to be much harder

to distinguish from genuine SNs. Spelling errors in SNs have also been left uncorrected; their high frequency is a characteristic feature in Polish SNs (Zaśko-Zielińska, 2013). SNs have been divided by the clustering algorithm into two subgroups (the maximum number of clusters was limited to 4 for the whole corpus). One subgroup showed no emotional content while the other was emotionally charged. Emotions were recognised on the basis of the annotation of words in the LIWC dictionary.

Text analysis of the suicide discourse in literature and poetry has also been attempted. Stirman and Pennebaker (2001) treated word use as an indicator of the mental states of suicidal and non-suicidal poets. Mulholland and Quinn (2013) applied the LIWC tool and dictionary in the analysis of over 70 language dimensions: polarity, affect states, death, sexuality, tense, etc. The dimensions were recognised on the basis of the word annotations in LIWC. The annotation and processing were done for words, not for word senses. Mulholland and Quinn (2013) tried to classify lyricists as suicidal or non-suicidal by their work and their known life stories. The goal of this research was to predict the likelihood of a musician committing suicide. The 70.6% classification rate represents a 12.8% increase over the majority-class baseline in the collected training set.

In (Pestian et al., 2010) a special suicide ontology containing 19 different classes of emotions was prepared and then used to annotate suicide notes. After the feature selection process the final four emotion concepts remained: hopelessness - regret - sorrow - giving things away (that is not strictly speaking an emotion). The final classifier working on the four emotion concepts and also on 42 specific words (among them prepositions, proper names, auxiliary verbs, the words good and love) outperformed mental health professionals in discerning elicited notes from real suicide notes. Following this publication, a special suicide note corpus annotated with 16 emotions was prepared in 2011.

3 Language Data

As the main source for training and testing we used the Polish Corpus of Suicide Notes (PCSN) (Zaśko-Zielińska, 2013). The PCSN is one of very few such resources in the world, e.g. it is significantly larger than the similar collection discussed by (Matykiewicz et al., 2009). It includes 1,244

genuine SNs that have been scanned and manually transcribed. Each SN was manually corrected and linguistically annotated on several levels, including selected semantic and pragmatic phenomena (Zaśko-Zielińska, 2013). The correction was necessary, as the originals include many errors or *ad hoc* abbreviations, that would be very difficult for automated processing. The annotation is stored in a TEI-based format (Marcińczuk et al., 2011) with corrected version in a separate layer. PCSN includes also a subcorpus of 334 counterfeited SNs (elicited). They were created by volunteers who were asked to imitate a real SN for imaginary person whose characteristic had been provided at the beginning of the experiment. The characteristics were randomly generated in a way following the distribution observed among the authors of PCSN genuine notes (the information is stored in the meta-data). Most volunteers were told that the notes written by them would be used 'to deceive' the computer program.

The genuine notes have varied length, but most of them are relatively short (around several sentences). Almost all of them were handwritten, while the counterfeited are all handwritten. The genuine notes include a lot of language errors, while the counterfeited are written in almost correct way. In such a situation, the errors are very clear signal for the genuine notes, that is why we used the corrected version and we used the layer of the corrected versions as the basis for the experiments. Thus the task was much more difficult. It is not clear if the same practice was implemented, e.g. in (Pestian et al., 2008).

As there is imbalance between the numbers of genuine and counterfeited SNs in PCSN, and the counterfeited SNs represent a specific genre, we have collected from Internet fora 2,200 letter-like texts. They represent a wide range of topics, but a have a form of a personal letter. In addition, we have randomly selected 1,000 Wikipedia articles, as examples of non-letters. All these collected texts were treated as negative examples during the experiments.

4 Descriptive Features

In a search for linguistic markers of SNs, we tested a rich set of features of the two main groups: *word-based* and *sense-based*. The first include lemmas (i.e. basic morphological forms), their annotations, derivation types and classes of Proper Names. The second group is based on word senses described in plWordNet (Piasecki et al., 2009) as synsets, their different generalisations, linguistic domains for synsets (Fellbaum, 1998; Piasecki et al., 2009) and the existing mappings of plWordNet onto SUMO ontology (Pease and Fellbaum, 2010; Pease, 2011).

Texts were pre-processed by *WCRFT* – a morpho-syntactic tagger for Polish (Radziszewski, 2013), *Liner2* – a named-entity recogniser (Marcińczuk et al., 2013) and *WoSeDon* – a prototype Word Sense Disambiguation tool (Kędzia et al., 2015) in a version that was based on plWordNet 2.2 (Maziarz et al., 2013). SNs were represented by such features as word lemmas, punctuation, text length, sentence length, grammatical classes of words, proper names and their classes.

4.1 Lexical and syntactic features

The set of word-based features on words and their annotations encompasses the frequencies of:

lemma – basic morphological forms from the tagger,

punctuation – punctuation marks,

big.letter – words started with a big letter,

gram.class – grammatical classes from the tagger,

verb12 – verbs in the 1st or 2nd person,

bigrams – bigrams of grammatical classes,

diminutive – diminutive forms identified on the basis of information from plWordNet,

augmentative – a similar feature to the above one,

PN.class – PN recognised by Liner2 as representing: first and last names, roads, cities and countries.

The feature *verb12* was intended to signal text of personal nature. Diminutives and augmentatives were expected to signal an emotional character, and PNs were assumed to appear more frequently in more concrete texts.

Some experiments replaced lemmas with word senses represented as plWordNet synonym set (synset) identifiers, assigned to words in the SNs by the WSD tool.

4.2 Semantically motivated features

The first group includes several features expressing clear semantic information, but in the case of the second group we use plWordNet as a basis. To compute features from the second group, words are mapped to plWordNet 2.2 synsets by WoSeDon. Its accuracy is limited and in practice it reaches about 75% on running text (the reported is lower) (Kędzia et al., 2015), but we assumed that the errors would not significantly influence the result. We used the following semantic features:

synsets – synsets from plWordNet 2.2,

hypernyms5 – all synsets on the hypernymic path up to five levels from the synset of the given word,

wn.domains – WordNet Domains (Bentivogli et al., 2004) assingned to synsets via mapping: plWordNet – Princeton WordNet (Fellbaum, 1998),

sumo – the first SUMO concepts accessible from the synset of the given word,

synset.hyp – hypernyms that are two levels above the word synset,

domain – linguistic domains of synsets,

verb.emo – verb lemmas described as expressing emotions in (Zaśko-Zielińska, 2013) on the basis of the analysis of plWordNet hypernymy structure,

noun.emo – in a similar way to the above one,

adj.emo – as above.

Synsets (as lemmas), can be too specific for particular SNs, and due to the limited number of SNs in the corpus they can fail in supporting the generalisation of the classifier. That is why we were looking into different ways of mapping synsets into classes defined by hypernyms, domains or SUMO concepts. However, the most sophisticated way of generalisation is described in the next section.

4.3 Class-related Sense Dictionaries

We aim at generalising particular words to dictionaries of senses that are characteristic for different types of contexts or texts. The underlying hypothesis of this approach is that generalisation of

specific words in a subset of documents from corpus allows to locate synsets in WordNet, for which we can reconstruct dictionaries, which describe the observed phenomenon and allows to distinguish between different types of words observed in the same set of documents. We adapted the algorithm presented in (Kocoń and Marcińczuk, 2016) for the purpose of selecting a subset of WordNet synsets which contain most specific words for each class of SN to improve classification of SNs. Algorithm 1 presents the dictionary generation.

On the basis of this method, we have generated dictionaries for three classes of texts: genuine SNs, counterfeited SNs and other texts. The dictionaries were generated from the held-out (tuning) subset. We calculated the frequency of synsets from a given dictionary as a feature. The use of the dictionary occurrence features is marked as *dictionary* in the description of the experiments.

5 Experiments and Results

The expanded PCSN was randomly divided into 10 parts. One of them was used for feature selection and generation of the class-related sense dictionaries. The rest was used for 10-fold cross validation. After the preliminary pre-experiments, we decided to use SVM classifier from the LIBSVM library (Chang and Lin, 2011) and the RBF kernel. During experiments we used different weighting methods. Three of them were tested in the final experiments: Pointwise Mutual Information, its version called Mutual Information in (Lin, 1998) and *tf* weighting (i.e. normalisation by the most frequent lemma/synset). Several other transformations did not bring improvement. All features less frequent than the threshold $f = 20$ and occurring in a smaller number of texts than $d = 5$ were filtered out. The feature values were scaled to the range $[0, 1]$ on the input to the SVM classifier.

We tested a large number of feature combinations, the best are presented in Table 1. They differ in the number of features selected by the InfoGain method on the held-out set, weighting method and the feature set used:

AnnLemmas = *lemmas*, *punctuation*, *gram.class*, *verb12*, *PN.class* and *bigrams*,

AnLem+Deriv = **AnnLemmas** plus *big.letter*, *diminutive* and *augmentative*,

NonPerLem = **AnLem+Deriv** minus *verb12*,

Algorithm 1 Construction of the class-related sense dictionaries for single class

Require:
 1: $G =< V, A >$ – WordNet as directed graph, where nodes V are synsets (sets of synonyms) and edges $A \in V \times V$ are hypernym relations;
 2: $\vec{d}$ – corpus as vector of words;
 3: t – semantic class (e.g. *genuine*);

Ensure:
 4: $\mathfrak{P}$ – dictionary of the *greatest positive correlations*;
 5: $\mathfrak{M}$ – dictionary of the *lowest negative correlations*;

 6: $updateGraph(G)$ — each synset $v \in V$ is extended with its hyponyms' lemmas;
 7: $classVector(\vec{d}, t)$ — construction of such a vector $\vec{w}$, where $|\vec{w}| = |\vec{d}|$ and $\vec{w}_n = 1$ if word $\vec{d}_n$ belongs to document classified as t, 0 otherwise;
 8: $synsetVector(\vec{d}, V)$ — for each $v \in V$ such a vector $\vec{a}^v$ is constructed, where $|\vec{a}^v| = |\vec{d}|$ and $\vec{a}^v_n = 1$ if $\vec{d}_n \in v$, 0 otherwise;
 9: $pearsonCorrelations(\vec{w}, \vec{a}, V)$ — for each $v \in V$ a Pearson correlation value is determined: $P^v = pearson(\vec{w}, \vec{a}^v)$;
 10: $bestNodes(V, P, p)$ — creating such synset $v \in V$ collections $\mathcal{P}$ and $\mathcal{M}$, for which P^v was the greatest ($\mathcal{P} \subseteq V$) and the lowest ($\mathcal{M} \subseteq V$) in each *hyponym branch*. Selection of the best nodes is dependent on parameter p, which specifies the minimal absolute value of Pearson correlation P^v to add v to $\mathcal{M}$ or $\mathcal{P}$. In experiments we used $p = 0.001$. for each pair $(v_i, v_j) \in \mathcal{M} \times \mathcal{M} \wedge i \neq j$ there is no path in G between (v_i, v_j), which means, that v_i, v_j cannot be in the same *hyponym branch*. The same applies to $\mathcal{P}$.
 11: $bestNodesSubsets(\mathcal{M}, \mathcal{P}, \vec{w})$ — this two-step method joins best nodes and builds two subsets: $\mathfrak{M} \subseteq \mathcal{M}$ and $\mathfrak{P} \subseteq \mathcal{P}$. In the first step a subset $\mathfrak{P}$ is constructed iteratively. In each iteration, the method searches for such element $e \in \mathcal{P}$, for which Pearson correlation $pearson(\vec{\omega}, \vec{w})$ is the greatest after the vector $\vec{\omega}$ is created ($|\vec{\omega}| = |\vec{d}|$ and $\vec{\omega}_n = 1$ if $\vec{d}_n \in \mathfrak{P} \cup \{e\}$, 0 otherwise). Next, $\mathfrak{P} = \mathfrak{P} \cup \{e\}, \mathcal{P} = \mathcal{P} \setminus \{e\}$ and a procedure is repeated until there is no Pearson correlation gain or $\mathcal{P} = \emptyset$. The second step looks similar. In each iteration the method searches for such element $e \in \mathcal{M}$, for which Pearson correlation $pearson(\vec{\omega}, \vec{w})$ is the greatest after the vector $\vec{\omega}$ is created ($|\vec{\omega}| = |\vec{d}|$ and $\vec{\omega}_n = 1$ if $\vec{d}_n \in \mathfrak{P}$ **and** $\vec{d}_n \notin \mathfrak{M} \cup \{e\}$, 0 otherwise). Next, $\mathfrak{M} = \mathfrak{M} \cup \{e\}, \mathcal{M} = \mathcal{M} \setminus \{e\}$ and a procedure is repeated until there is no Pearson correlation gain or $\mathcal{M} = \emptyset$.

Synsets = AnLem+Deriv minus *lemmas* plus *synsets, hypernyms5, wn.domains* and *sumo*,

GenSyn+Dom = AnLem+Deriv minus *lemmas* plus *synset.hyp, domain, verb.emo, noun.emo, adj.emo* and *sumo*,

Dom+SUMO = GenSyn+Dom minus *synset.hyp*,

SenseDict = GenSyn+Dom plus *dictionary*.

The first three vectors, namely **AnnLemmas**, **AnLem+Deriv** and **NonPerLem** do not refer to word senses and do not require pre-processing based on WSD. The basic **AnnLemmas** describe lemmas and punctuation occurring in texts with an intention to identify lemmas characteristic for genuine SNs. In addition bigrams of the grammatical classes provide some hints on syntactic structures, PNs show that text is more concrete and *verbs12* reveals the personal elements and instructions included in the text. **AnLem+Deriv** adds aspects of informal, emotional descriptions (positive and negative). In **NonPerLem** we wanted to find out what is the influence of the *verb12* feature.

Because we expected that words can be quite specific and accidental due to the limited set of documents, in the next group of vectors we tried to map the documents on the semantic space and open possibilities for different kinds of generalisations on the basis of the very large structure of plWordNet and SUMO linked together. The **Synsets** feature vector was the first attempt, in which words were exchanged by synsets and we traced paths across all synsets up several levels the hypernymy structure aiming at expanding the

Exp.	Weight.	Feat.	Acc	F	PosP	NegP	R	Spec	CounterP
Word-based									
AnnLemmas	–	500	88.34	77.78	79.51	91.39	76.13	92.81	69.67
AnnLemmas	PMI	500	93.28	87.14	89.35	94.62	85.04	85.04	73.44
AnnLemmas	MI	500	93.62	87.54	**91.47**	94.32	83.94	**97.15**	74.92
AnnLemmas	MI	1000	91.81	84.91	89.61	92.57	80.68	96.26	76.08
AnLem+Deriv	tf	604	**94.27**	**89.25**	90.40	95.66	88.13	96.54	**77.78**
NonPerLem	tf	603	93.78	88.46	88.56	**95.70**	**88.36**	95.78	74.36
Sense-based									
Synsets	MI	500	92.59	85.70	86.49	95.30	84.93	94.69	66.41
Synsets	MI	2000	92.13	85.34	86.75	94.06	83.95	95.20	70.00
GenSyn+Dom	tf	2000	93.78	88.46	88.56	95.70	88.36	95.78	70.09
Dom+SUMO	tf	1383	94.15	89.07	89.79	95.73	88.36	96.29	73.05
Domains	tf	653	94.33	89.38	**90.42**	95.74	88.36	**96.54**	**76.92**
GenSyn+Dom	tf	500	93.90	88.63	89.15	95.63	88.13	96.04	73.50
SenseDict	tf	2000	**94.64**	**90.06**	90.16	**96.29**	**89.95**	96.37	74.36

Table 1: Results of the classification of Suicide Notes on the basis of different feature vectors (Feat. – the number of the selected features, $Acc = \frac{TP+TN}{TP+FP+TN+FN}$, $PosP = \frac{TP}{TP+FP}$, $NegP = \frac{TN}{TN+FN}$, $R = \frac{TP}{TP+FN}$ – recall, $Spec = \frac{TN}{TN+FP}$, $CounterP$ – the precision in the subset of counterfeited SNs).

description by more general synsets (i.e. lexical meaning), too. As a means of generalisation of the description. In addition we added the mapping to SUMO concepts (that seemed to work well), as an even further generalisation, and WordNet Domains (that introduced too much noise[1]). In the next group of semantic vectors: **GenSyn+Dom**, **Dom+SUMO** and **SenseDict** synsets have been exchanged with their medium-grained generalisations, i.e. every synset was mapped onto a hypernym two levels up and without adding all synsets from the path as it was done in **Synsets**. Moreover, we also used wordnet linguistic domains that were introduced to support wordnet editors (Fellbaum, 1998), but appeared to be a useful way of grouping senses in at least several applications. **Dom+SUMO** do not include synset-based features, but instead mappings to SUMO, while **SenseDict** extends the synset-based vector with the proposed way of extracting domain-related sense dictionaries.

The results were evaluated according to the 10-fold evaluation scheme performed on the training-test set. The average values of the several standard measures across the folds are given in Table 1. The F measure is calculated from the precision $PosP$ and the recall R, which shows how many genuine SNs were recognised. As the counterfeited SNs should be all filtered out by an ideal classifier, we

have introduced a separate precision measure for this subset, namely $CounterP$.

In Table 1 we can see that all proposed models present very good performance in general, that is superior to the results achieved so far in the literature for similar tasks. $CounterP$ is much lower, but still much above 50% baseline and this is the most difficult subtask. Moreover in spite of the fact that the size of the set of counterfeited SNs much smaller than the others set, $CounterP$ is still larger than the results reported in the literature.

The results of the first experiment are slightly lower due to the lack of weighting. Word-based models and synsets-based models express similar performance if some mechanisms for generalisation are introduced to the latter, e.g. the simpler **Synsets** model which uses many more specific synsets produced lower results. In the same time **Domains** model that does not refer to synsets and **SenseDict** that utilises classes of word senses achieved higher performance than word-based model. The difference is on the margin of the statistical significance, e.g. in the case of **SenseDict** the difference is on the 95% level of trust. However, we can conclude that looking for the ways of wordnet-based generalisation of the representation is worth attention. The difference between **NonPerLem** and **AnLem+Deriv** shows that the influence of *verb12*, that was meant to represent personal elements in the note, is not clear. On the one hand **NonPerLem** has the best value of $CounterP$, but on the other hand the feature

[1]WordNet Domains were extracted automatically from a large English corpus and next transferred from Princeton WordNet to plWordNet via the manually created inter-lingual mapping – too many places in which noise could appear.

verb12 was selected as a significant one during feature selection for models in which it was included to.

The good results obtained, especially with the classification based on semantic features suggest that the linguistic content of SNs is a strong factor separating them out from other types of writing including non-personal and personal texts (namely letters). Moreover, the linguistic features of SNs make them also different from the counterfeited SNs that were written by humans with intention to deceive 'a computer program'. So we can expect that subjects did all their best during experiments, but still the language used by them express enough differences to be captured by our classifiers.

The feature vectors that produced the best results give some insights into the character of the linguistic differences between true SNs and the other types of writing. In order to take a closer look we have examined the ranking of features selected for the **SenseDict** vector on the basis of the InfoGain algorithm and the held-out set. The top 45 features are presented in Tab. 2. Most of the labels used to name the features are explained in the caption. However the names of the specific plWordNet synsets were too long to fit them into the tables:

synHyp:Group – synHyp:{*grupa* 4 'a group', *zbiór* 1 'a set'}

synHyp:Property – synHyp:{ *właściwość* 1 'property', *przymiot* 1 'attribute' *cecha* 1 'characteristic feature',{ *własność* 2 'property', *atrybut* 1 'attribute'}

sumo:SbjAssessmentAtr – sumo:subsumed by SubjectiveAssessmentAttribute

synHyp:makingRelMag – synHyp:{ [non lexicalised] *wykonywanie czynności religijnych badz magicznych* 1 'performing religious or magical acts'}

synHyp:going away – synHyp:{ *oddalanie si?* 1 'going away' or 'passing away' }

synHyp:ManSocialRel – synHyp:{ [non-lexicalised] *człowiek ze względu na relacje społeczne* 1 'a man distinguished by its social relationships' }

synHyp:State – synHyp:{ *stan* 1 ' a state'}

synHyp:GerDynVerb – synHyp:{ [non-lexicalised, a top synset for a class of gerund nouns] *GERUNDIUM OD CZASOWNIKA DYNAMICZNEGO NIEZMIENNOSTANOWEGO* 1 'a gerund noun derived from a dynamic verb not imposing a change of state'}

Synset dictionaries constructed for general texts (*tex dict.*), as well as genuine and counterfeited SNs are among the top features in Tab. 2. A very high position of *verbs12* seem to reveal the personal character of SNs. The significance of different punctuation symbols is specific for SNs. The general class of punctuations (*lexClass:interp*) is high on the list, but also many bigrams with punctuations (e.g. bigrams:adj+interp), and individual symbols (e.g. *interp:comma*) are close to the top. According to the linguistic analysis, imperative forms of verbs are frequent in genuine SNs, and this is confirmed by *lexClass:impt* representing this specific grammatical class. In the top of the feature ranking, we can also notice several specific semantic features: top hypernyms for different senses referring to groups of people (*synHyp:Group*) including 'a family', for all kinds of situations (*synsetHyp:GERUNDIUM*) but also specific situations of religious acts (e.g. praying) and passing away (*synHyp:going away*). The synset *synHyp:ManSocialRel* dominates many synsets representing social roles of people and this may be caused by frequent referring by authors of SNs to family members or people related to them by naming social roles of those people. The concept *sumo:SbjAssessmentAtr* subsumes many synsets describing man's character – SNs are full of positive and negative description of people. Finally, the specific grammatical class of *aglt* signals more frequent use of the subjunctive mood.

6 Conclusions

The obtained results show that there are fundamental differences between genuine SNs and counterfeited SNs. The differences are even striking in relation to other types of texts. It is worth emphasising that we compared the transcribed versions of SNs not taking into account different types of errors that occur very often in them, the length of the letters, layout of the letters etc. The analysis was intentionally focused only on linguistic properties and the selected feature vectors re-

No	Feature	No	Feature	No	Feature
1	text dict.	16	counterfeited dict.	31	bigrams:num+subst
2	lexClass:subst	17	bigrams:prep+subst	32	bigrams:interp+subst
3	bigrams:interp+empty	18	lexClass:impt	33	PN:country
4	lexClass:interp	19	bigrams:interp+interp	34	domain:zdarz
5	bigrams:adj+interp	20	lexClass:ger	35	synsetHyp:GERUNDIUM
6	genuine dict.	21	interp:question	36	sumo:SbjAssessmentAtr
7	verb12	22	bigrams:adj+subst	37	synHyp:makingRelMag
8	bigrams:subst+interp	23	interp:hyphen	38	synHyp:going away
9	lexClass:ppron12	24	bigrams:subst+subst	39	interp:dash
10	interp:comma	25	interp:fullstop	40	synHyp:ManSocialRel
11	lexClass:noun	26	bigrams:interp+adj	41	synHyp:State
12	lexClass:prep	27	bigrams:subst+ppas	42	synHyp:GerDynVerb
13	bigrams:subst+adj	28	synHyp:Group	43	bigrams:adj+prep
14	domain:rel	29	synHyp:Property	44	lexClass:aglt
15	lexClass:adj	30	bigrams:ppas+prep	45	bigrams:ppron12+praet

Table 2: Characteristic features selected for classifier based on **SenseDict** vector (where *dict.* – a domain dictionary of synsets, *domain:rel* – the domain of relative adjectives, and *domain:zdarz* – domain of event verbs, *lexClass* – a grammatical class. (with *aglt* – aglutinative participle used, e.g., for subjunctive mood, *ger* – gerund, *impt* – imperative verb form, *interp* – punctuation symbol, *num* – numeral, *ppas* – perfective adjectival participle, *ppron12* – personal pronoun of 1st or 2nd person, *praet* – past verb from, but also used for compound future time, *prep* – preposition, *subst* – noun), *verb12* – verbs in 1st or 2nd person.

vealed many features that are characteristic for SNs. In many cases they corresponded to the features identified manually in (Zaśko-Zielińska, 2013).

The models based on synsets are only slightly better than those on words, but the former seem to offer natural ways of generalising the description. The applied method of the sense dictionary construction appeared to be the best way of improving the model. The applied WSD was of a limited accuracy, but it has still some room for improvement.

References

Luisa Bentivogli, Pamela Forner, Bernardo Magnini, and Emanuele Pianta. 2004. Revising wordnet domains hierarchy: Semantics, coverage, and balancing. In *COLING 2004 Workshop on Multilingual Linguistic Resources Geneva, Switzerland, August 28*. pages 101–108.

Chih-Chung Chang and Chih-Jen Lin. 2011. LIB-SVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Christiane Fellbaum, editor. 1998. *WordNet – An Electronic Lexical Database*. The MIT Press.

J. M. Gomez. 2014. Language technologies for suicide prevention in social media. In *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC 2014)*. pages 21–29.

E.A. Holmes, C. Crane, M. J.V. Fennell, and Williams J.M.G. 2007. Imagery about suicide in depression – "flash-forwards"? *Journal of Behavior Therapy and Experimental Psychiatry* 38:423–434.

Paweł Kędzia, Maciej Piasecki, and Marlena J. Orlińska. 2015. Word sense disambiguation based on large scale Polish clarin heterogeneous lexical resources. *Cognitive Studies* 14(To appear).

Jan Kocoń and Michał Marcińczuk. 2016. Generating of Events Dictionaries from Polish WordNet for the Recognition of Events in Polish Documents. In *Text, Speech and Dialogue, Proceedings of the 19th International Conference TSD 2016*. Springer, Brno, Czech Republic, volume 9924 of *Lecture Notes in Artificial Intelligence*.

Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In *Proceedings of the Joint Conference of the International Committee on Computational Linguistics*. ACL, pages 768–774.

M. Marcińczuk, J. Kocoń, and M. Janicki. 2013. Liner2 – a customizable framework for proper names recognition for Polish. In *Intelligent Tools for Building a Scientific Information Platform*, Springer, volume 467 of *Studies in Computational Intelligence*, pages 231–253.

Michal Marcińczuk, Monika Zaśko-Zielińska, and Maciej Piasecki. 2011. Structure annotation in the Polish corpus of suicide notes. In Ivan Habernal and Václav Matousek, editors, *Text, Speech and Dialogue - 14th International Conference, TSD 2011,*

Pilsen, Czech Republic, September 1-5, 2011. Proceedings. Springer, volume 6836 of *Lecture Notes in Computer Science*, pages 419–426.

P. Matykiewicz, W. Duch, and Pestian. J. 2009. Clustering semantic spaces of suicide notes and newsgroups articles. In *Proceedings of the Workshop on BioNLP*. pages 179–184.

Marek Maziarz, Maciej Piasecki, Ewa Rudnicka, and Stan Szpakowicz. 2013. Beyond the Transfer-and-Merge Wordnet Construction: plWordNet and a Comparison with WordNet. In G. Angelova, K. Bontcheva, and R. Mitkov, editors, *Proceedings of International Conference on Recent Advances in Natural Language Processing*. Incoma Ltd., Hissar, Bulgaria.

M. Mulholland and J. Quinn. 2013. Suicidal tendencies: The automatic classification of suicidal and non- suicidal lyricists using nlp. In *International Joint Conference on Natural Language Processing*. pages 680–684.

Adam Pease. 2011. *Ontology - A Practical Guide*. Articulate Software Press, Angwin, CA.

Adam Pease and Christiane Fellbaum. 2010. Formal ontology as interlingua: the SUMO and WordNet linking project and Global WordNet. In Chu-Ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Oltramari, and Laurent Prévot, editors, *Ontology and the Lexicon. A Natural Languge Processing Perspective*, Cambridge University Press, Studies in Natural Languge Processing.

J. W. Pennebaker, M. E. Francis, and R. J. Booth. 2001. *Linguistic Inquiry and Word Count: LIWC*. Lawrence Erlbaum Associates, Mahwah.

J.W. Pennebaker and C. K. Chung. 2011. Expressive writing: Connections to physical and mental health. In H. S. Friedman, editor, *The Oxford Handbook of Health Psychology*, Oxford University Press, pages 417–437.

J. Pestian, H. Nasrallah, P. Matykiewicz, A. Bennett, and A. Leenaars. 2010. Suicide note classification using natural language processing. *Biomed Inform Insights* 3:19–28.

John P. Pestian, Pawel Matykiewicz, and Jacqueline Grupp-Phelan. 2008. Using natural language processing to classify suicide notes. In *BioNLP 2008: Current Trends in Biomedical Natural Language Processing*. pages 96–97.

Maciej Piasecki, Stanisław Szpakowicz, and Bartosz Broda. 2009. *A Wordnet from the Ground Up*. Wrocław University of Technology Press. http://www.eecs.uottawa.ca/~szpak/pub/A_Wordnet_from_the_Ground_Up.zip.

Adam Radziszewski. 2013. A tiered CRF tagger for Polish. In *Intelligent Tools for Building a Scientific Information Platform*, Springer, volume 467 of *Studies in Computational Intelligence*, pages 215–230.

E.S. Shneidman and N.L. Farberow, editors. 1957. *Clues to Suicide*. Blakiston Division, New York.

S. W. Stirman and J.W. Pennebaker. 2001. Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine* 63:517–522.

World Health Organisation. 2014. Preventing suicide: A global imperative. Technical report, World Health Organization.

Monika Zaśko-Zielińska. 2013. *Listy pożegnalne: w poszukiwaniu lingwistycznych wyznaczników autentyczności tekstu*. Wydawnictwo Quaestio, Wrocław.