

Personalized linguistic summaries in smartphone-based monitoring of bipolar disorder patients

Katarzyna Kaczmarek-Majer^a and Olgierd Hryniewicz^a and Monika Dominiak^{b,c} and Łukasz Świącicki^b

^aDepartment of Stochastic Methods, Systems Research Institute, Polish Academy of Sciences,
Newelska 6, 01-447 Warsaw, Poland,

k.kaczmarek@ibspan.waw.pl, olgierd.hryniewicz@ibspan.waw.pl

^b Department of Affective Disorders, Institute of Psychiatry and Neurology,
Sobieskiego 9, 02-957 Warsaw, Poland
swiecicki@ipin.edu.pl, mdominia@wp.pl

^c Department of Pharmacology, Institute of Psychiatry and Neurology
Sobieskiego 9, 02-957 Warsaw, Poland

Abstract

Personalized linguistic summaries are developed with the use of protoforms in the sense of Yager and Kacprzyk. We discuss the construction and usefulness of such patient-dependent and disease-state-dependent linguistic summaries that may be exemplified as *Most outgoing calls in mania state (disease period) are short compared to the calls recorded in the euthymia state (healthy period)*. Such linguistic summaries may become important features in the smartphone-based monitoring of the bipolar disorder patients and inform about the detected change in patient's state. The main advantage of the personalized linguistic summaries is their human-centricity. The performance of the proposed approach is illustrated with examples based on the real-life data collected within the observational study.

Keywords: Linguistic summaries, Time series, Prediction, Bipolar disorder

1 Introduction

Bipolar disorder (BD) is a mental disorder that affects more than 2% of the world's population [8]. It is a chronic and recurrent disease with the highest rate of suicide of all the psychiatric disorders. It is characterized by manic episodes of elevated mood and overactivity, interspersed with periods of depression' [4]. Recent studies, e.g., [7], show that smartphone becomes an increasingly effective tool for monitoring the state of a BD patient and early detection of a state change. The objective data automatically collected using smartphones become valid markers of a mood state [9]. Kaczmarek-Majer et al. [17] recently showed that statistical process control is an adequate methodology to build patient-dependent models and

generate alarms when the patient's behaviour related to the smartphone usage changes.

For example, in Figure 1, the control chart of the smartphone-based data representing the daily mean duration of outgoing calls for an exemplary patient P is presented. Patient P was in healthy state (euthymia state) from 1st day to 20th day and in the disease state (depression) from 21st day.

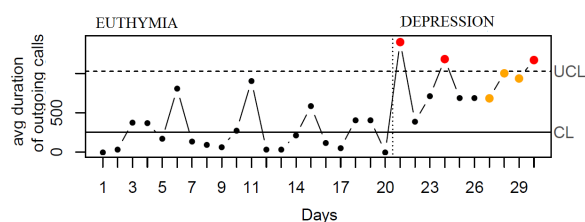


Figure 1: Control chart for the smartphone-based data representing the daily mean duration of outgoing calls for patient P

As depicted in Figure 1, the observation fell beyond the upper control limit (UCL) and there were alarms generated on day 21st (1st day of the assessed depression), 24th and 30th. This was a positive aspect that an alarm was generated already on day 21st, because the patient's state indeed changed to depression (as resulted within the clinical assessment).

However, efficient human-machine communication with the psychiatrist about an alarming situation caused by the smartphone-based data, for example via short text message, is challenging. Informing a psychiatrist about an absolute number of outgoing calls or their mean duration is not a solution because each patient has his own habits about using smartphone. Although, in the phase of disease (depression or mania) the patient's behavior changes, and it differs from the state of euthymia, the precise direction of a change depends on a patient and the severity of the depressive and manic symptoms. Nonetheless, it is very impor-

tant to informatively describe the change related to the monitored smartphone-based data when communicating with a psychiatrist.

Within this research, we build patient-dependent and disease-state-dependent linguistic summaries and call them personalized linguistic summaries. The following types of linguistic summaries are considered:

- *Most outgoing calls of patient A in disease state (e.g., depression) are long.*
- *Most outgoing calls of all patients in disease state (e.g., depression) are long.*

Linguistic summaries are easy for interpretation, and thus, are expected to facilitate the communication with a psychiatrist about an alarming situation related to the smartphone-based data. As observed in practice, linguistic summaries constructed with general (patient-independent) linguistic variables fail to adequately describe the smartphone-based data, therefore, we pursue a personalized approach. Within this research, the fuzzy numbers are constructed basing on the historic smartphone-based data of an individual patient in a healthy state (euthymia). Furthermore, we provide preliminary results assessing whether there are differences between degrees of truth of personalized linguistic summaries about duration of outgoing calls in various BD phases.

2 Related work

Linguistic summarization is one of the descriptive techniques of knowledge discovery. Many studies confirmed that summarization techniques are able to extract knowledge from large databases, and by doing this, to improve the comprehension of it or to enable easier processing of large volumes of data streams [20]. It was originally based on the Yager's concept of protoform [26] and Zadeh's fuzzy quantifiers [27] further developed by Kacprzyk [13, 15]. Its first domain of application was transactional databases, however, it quickly encompassed time series (and sequential data) for which linguistic summaries may express the general trends, e.g., *Among all increasing segments, majority is long.* cf. Kacprzyk et al. [12]. One of their main advantage is human-consistency. In [18], Lesot et al. investigate the interpretability of fuzzy linguistic summaries, both at the sentence level and at the group of summaries level.

Boran, Akay and Yager [5] provide the overview of recent methods for linguistic summarization with fuzzy sets. Recently, also Marin et al. [19] review the main contributions in the fields of Natural Language Generation and Fuzzy Sets and Systems, and provide a gen-

eral approach for the concepts and processes related to the generation of linguistic descriptions of time series. Marin's approach consists of two main tasks, namely, a knowledge extraction task, which can be seen as a knowledge discovery in databases procedure, and a linguistic expression process. The linguistic descriptions, see also [24, 2], are much complex in terms of semantic and lexico-grammar structures, e.g., *Before the knee lesion, the gait quality is high because the gait symmetry is medium and the gait homogeneity is high* cf. Alvarez-Alvarez and Trivino [3].

In the recent paper by Kacprzyk et al. [14], the authors claim that linguistic data summaries in Yager's sense can be considered an »ultimately human consistent« form of human-centric aggregation. This paper extends our previous works devoted to the statistical process control that aimed at the monitoring of the autocorrelated health-related processes [10, 11, 17] and works dedicated to the incorporation of linguistic summaries into time series forecasting [16].

Linguistic summaries have been proven successful in various healthcare-related context, e.g., in [25, 6, 1, 23]. For example, in [25], Wilbik et al. focus on potential changes in time and introduce modification of summaries into the linguistic medoid prototypes, that may be exemplified with the following sentence: *Mrs. Wilbowski is much better this week than she was last week.* Moyse et al. [21] focus on local periodic components of a time series and propose summaries like *Approximately from March to June, the series is highly periodic with a period of exactly 2 weeks.* For further extensions of linguistic summaries see also the evaluative linguistic expressions introduced by Novak [22]. Within this research, we focus on the patient-dependent and disease-state-dependent linguistic summaries.

3 Linguistic summarization. Basic definitions.

Linguistic summaries describe with quasi natural language the general facts about evolution of a time series. To formally describe the linguistically quantified sentences that summarize a time series, we adapt the classic calculus of linguistically quantified propositions and the concept of the protoform in the sense of Yager [26] further developed by Kacprzyk et al. [15].

Let $O = \{o_1, o_2, \dots, o_b\}$ denote a finite set of objects (e.g., line segments of time series) in a considered domain. The properties of objects are measured by observables and are called attributes. Let $A = \{a_1, a_2, \dots, a_r\}$ denote a finite set of attributes (e.g., dynamics of change, duration), and $S = \{s_1, s_2, \dots, s_l\}$ is a finite set of labels for attributes (e.g., 'increasing',

‘short’). Quantifiers, qualifiers and summarizers are modeled as linguistic variables using fuzzy numbers.

Definition 1. Short protoform linguistic summary [26]

Short protoform linguistic summary (LS_{short}) takes the form: ‘Among all objects, Q are P [T]’ where Q is the quantifier (the amount determination, e.g., ‘most’, ‘minority’); P is the summarizer (attribute together with an imprecise label, e.g., ‘increasing segment’) about objects $o \in O$ and T measures the quality of the summary (level of confidence) being a real number from $[0,1]$.

Linguistic summary based on the short protoform may be exemplified by the following sentence: *Among all outgoing calls, most are short* [0.9].

One of the first measures to evaluate the quality of linguistic summaries is the degree of truth (validity) (d_T) introduced by Zadeh [27].

Definition 4. Degree of truth [27]

Let $\mu_R, \mu_P, \mu_Q : Re \rightarrow [0,1]$ be the membership functions of fuzzy sets representing the qualifier R , summarizer P and quantifier Q , respectively, and \wedge is the t -norm. Degree of truth is defined as follows

$$d_T(LS_{short}) = \mu_Q\left(\frac{1}{n} \sum_{i=1}^n \mu_P(y_i)\right) \quad (1)$$

4 Personalized linguistic summaries

Patient-dependent and disease-state-dependent linguistic summaries are developed based on the protoform concept. However, contrary to the general linguistic summarization approach, the fuzzy numbers for the linguistic variables are constructed taking into account the historical data of an individual patient in a healthy state.

Definition 3. Personalized short protoform linguistic summary

*Personalized linguistic summary based on the short protoform takes the form ‘Among all objects **belonging to patient A**, Q are P [T]’ where Q is the quantifier (the amount determination, e.g., ‘most’, ‘minority’); P is the summarizer (attribute together with an imprecise label, e.g., ‘short call’) about objects $o \in O$ and the fuzzy set interpretation for the imprecise label is established according to the stable (healthy) period of patient A; and T measures the quality of the summary (level of confidence) being a real number from $[0,1]$.*

Personalized short protoforms linguistic summaries may be exemplified by the following sentences:

Among all outgoing calls of patient A, most calls recorded in disease state are short (compared to the

calls recorded in healthy state). [0.9]

Among all sent text messages of patient A, most messages in disease state are long [0.5]

The number of outgoing calls of patient A in disease state is high [0.8]

When considering applications in medicine, apart from human-centricity, another advantage of processing personalized linguistic summaries instead of actual numbers may be avoiding storage of sensitive data about patient’s historical records.

5 Numerical results

5.1 Observational study

The observational study included patients diagnosed with bipolar disorder (according to ICD-10 classification) and was conducted in the Department of Affective Disorders, Institute of Psychiatry and Neurology in Warsaw. The study aimed at the development of a dedicated smartphone-based diagnostic tool for detection and prediction of the early symptoms of a phase change (BDmon app). BDmon was installed on patients’s smartphone’s to work in the background and record both, daily statistics about calls and text messages and features about patient’s voice during phone calls. Within this paper, we present preliminary results about patient-dependent and patient-independent linguistic summaries that are related to one of the monitored parameters that is smartphone-based daily mean duration of outgoing calls.

The psychiatric assessment of study participant was assessed every two months or more frequently, if requested, by a patient or a doctor. The psychiatric assessment was based on the Hamilton Depression Rating Scale (HDRS-17), Young Mania Rating Scale (YMRS), Clinical General Impression Scale (CGI-BD) and Beck Depression Inventory (BDI). The mental state of study participants was also evaluated fortnightly via phone-based interactions with a clinician.

The smartphone-based data collection was performed during the everyday life of a patient. However, the number of labeled data is reduced only to days around the clinical assessment. The ground-truth for the analyses was assumed to be 7 days before the clinical visit and 2 days after [9]. Also, the sample is reduced to patients who were at least 10 days in the euthymic state and used the smartphone within this period. As a result, data of 23 patients were validated for further analysis and linguistic summarization.

5.2 Illustrative example: smartphone-based monitoring of a patient going from euthymia to mania

First, we illustrate the performance of the patient-dependent and patient-independent linguistic summaries related to the duration of outgoing calls for an exemplary patient that experienced both, the healthy state (euthymia) and the disease state (mania). We start from explaining the observed problem and presenting the descriptive statistic of the considered variable: smartphone-based daily mean duration of outgoing calls.

Descriptive statistics about duration of outgoing calls. In Figure 2, the daily mean time (duration) of outgoing calls in various BD phases is illustrated with boxplots and Tukey plots.

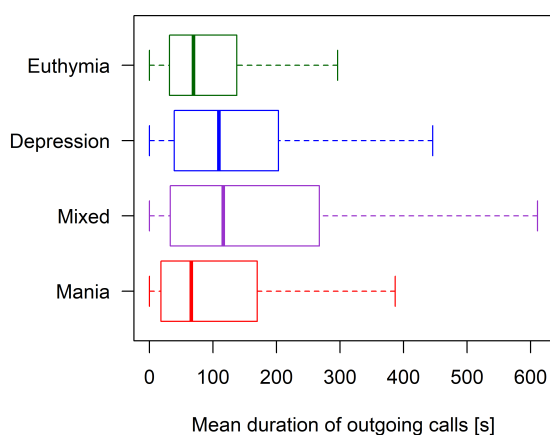


Figure 2: Boxplots of daily mean duration of outgoing calls for patients in various BD phases.

As observed, the mean duration of outgoing calls is rather similar for various patients in the four BD phases. We have further tested several hypotheses to check whether duration of outgoing calls can be considered as a valid marker for the assessment of the phase in BD. Tukey's test is a statistical test that verifies if the means of the 2 groups are significantly different from each other.

Tukey plots show differences (mean difference \pm SD of the difference) between values of the considered variable in two groups (e.g., patients in euthymia vs. patients in depression). The intervals, which do not cover zero denote significant difference and are plot in red. As observed in the Tukey plot in Figure 3, the daily mean duration of outgoing calls differs significantly for patients in euthymia state vs. patients in depression (p -value <0.001) and for patients in the mixed state vs. patients in euthymia (p -value <0.001). Tukey plots in-

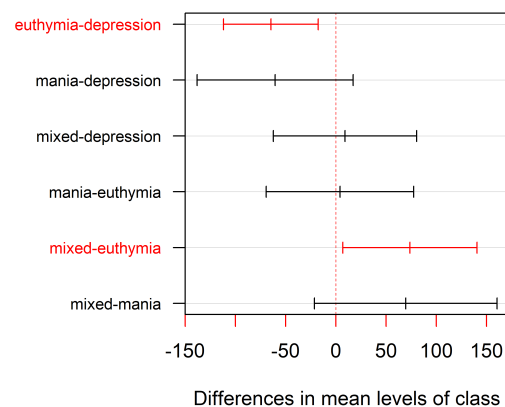


Figure 3: Tukey plot with daily mean duration of outgoing calls for patients in various BD phases

form also about the sign of the difference. For example, for the considered mean duration of outgoing calls, the euthymic patients make shorter calls than patients in depression or mixed state. Basing on these results, surprisingly, there is no difference in call duration for patients in mania vs. patients in euthymia.

Descriptive statistics for patient A. Now, we analyze duration of outgoing calls of an exemplary patient A – a female, living with family in a city, being a pensioner in her 50s. Patient A was considered euthymic during one clinical assessment. Then, the state of patient A was assessed as mania during another clinical assessment (4 BDI points, 3 points on HAMD scale, 11 points on YMS scale). In Figure 4, we illustrate with boxplots the daily mean duration of the outgoing calls for this patient A depending on the BD phase.

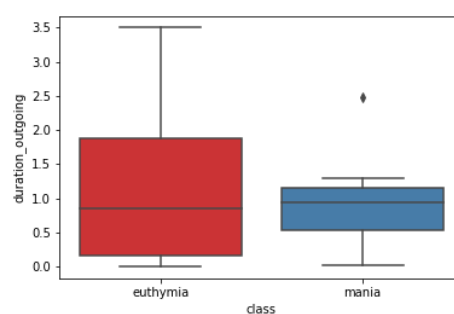


Figure 4: Boxplot of daily mean duration of outgoing calls (in minutes) for patient A in euthymia and mania

As presented in Figure 4, the median (1/2 quartile marked with the horizontal line) duration of outgoing calls for patient A in mania vs. median duration in euthymia are almost the same and amount to one minute. Similarly to the results depicted in Figure 3,

Attribute	f_1	f_2	f_3	f_4
<i>short duration</i>	a	a	$a + (b-a)/4$	$a + (b-a)/2$
<i>medium duration</i>	$a + (b-a)/4$	$a + (b-a)/3$	$a + (b-a)*2/3$	$a + (b-a)*3/4$
<i>long duration</i>	$a + (b-a)/2$	$a + (b-a)*3/4$	b	b

Table 1: Attributes of *duration of outgoing calls* modeled as trapezoidal fuzzy numbers $[f_1, f_2, f_3, f_4]$ where $a = \min(data)$ and $b = \max(data)$.

Linguistic summary	Patient-independent	Personalized
Most outgoing calls of Patient A in mania are short	1.0	1.0
Most outgoing calls of Patient A in euthymia are short	1.0	0.74

Table 2: Degree of truth of patient-independent linguistic summaries vs. personalized linguistic summaries (about exemplary patient A) related to the duration of outgoing calls.

there is almost no difference in distribution of data in the manic and the euthymic state.

Construction of linguistic variables. Before building the linguistic summaries, we construct fuzzy numbers for the attributes of linguistic variables. We model them as fuzzy trapezoidal numbers as defined in Table 1 depending on the minimum and maximum calculated from historic data.

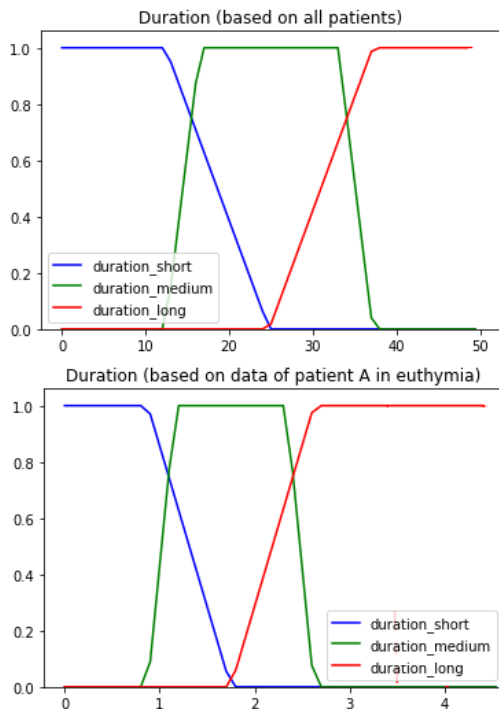


Figure 5: Linguistic variable *duration* (of the outgoing calls) (in minutes) and its interpretations based on (a) all data; (b) data for patient A in healthy state

First, we establish the fuzzy numbers basing on the

historic data of all patients (a and b are calculated basing on duration of outgoing calls of all patients). Secondly, we establish the linguistic variables basing on the data recorded only for an individual patient in a healthy phase (euthymia) (a and b are calculated basing on duration of outgoing calls of and individual patient in euthymia). In Figure 5., two considered ways of the construction of fuzzy sets representing the linguistic terms (*short*, *medium*, *long*) of the linguistic variable *duration of outgoing call* are illustrated.

As depicted in Figure 5., the interpretations of linguistic terms (*short*, *medium*, *long*) for the duration of outgoing calls is established based on (the statistics of) all historical data are different from the interpretations established based on data of a single patient in a healthy state. For example, 10-minute long outgoing call is short [1.0] according to (a) and is long [1.0] according to (b).

Linguistic summaries. Next, we evaluate the quality of linguistic summaries generated with two considered approaches to linguistic variables selection (patient-independent vs. patient-and-disease-state-dependent). In Table 2, the degree of truth of linguistic summaries of the form: *Most calls recorded in mania/euthymia are short* are compared.

As observed in Table 2, the patient-independent linguistic summaries: *Most outgoing calls in mania are short* and *Most outgoing calls in euthymia are short* are both completely true. This result is consistent with the outcome of Figure 2. However, this is not an expected outcome, because there tend to be differences between some patients in regards to the calling habits in mania state [9]. The personalized linguistic summary *Among all outgoing calls of patient A, most calls recorded in mania are short* is completely true (DoT=1.0), but the second considered personalized linguistic summary about the duration of calls

Linguistic summary description	Euthymia	Mania	Depression	Mixed
<i>Most outgoing calls are short</i>	0.00	1.00	1.00	0.39
<i>Most outgoing calls are medium</i>	1.00	0.83	0.67	1.00
<i>Most outgoing calls are long</i>	0.59	0.00	0.00	0.00

Table 3: Personalized linguistic summaries about duration of outgoing calls in various BD states.

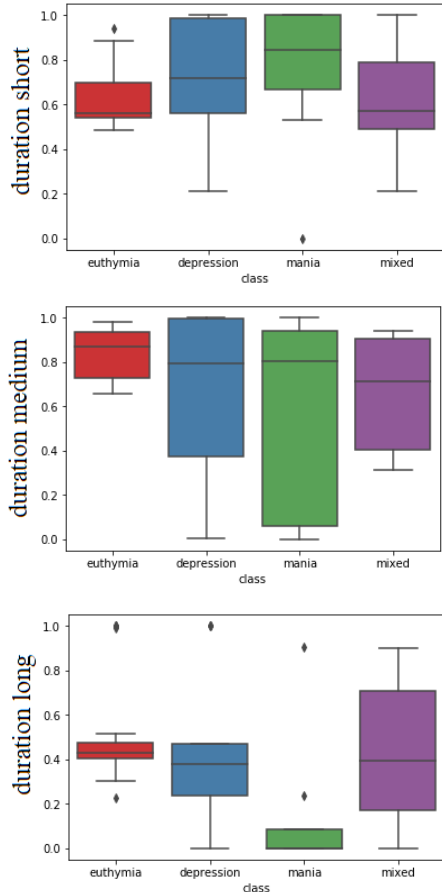


Figure 6: Boxplots of degree of truth of *short*; (b) *medium*; (c) *long duration of outgoing calls* for patients in various BD phases.

in euthymia is true only to some degree (DoT=0.74). We conclude that the personalized linguistic summary *Most outgoing calls of patient A are short* is more true in mania than in the euthymia state. In the next section, we present results of the personalized linguistic summarization about duration of outgoing calls using all validated data from the observational study.

5.3 Personalized linguistic summaries about duration of outgoing calls

The idea was to construct personalized (patient-and-disease-state-dependent) linguistic summaries and to

assess whether there are differences between degrees of truth of personalized linguistic summaries about duration of outgoing calls in various BD phases. First, we summarize the descriptive statistics about the calculated degrees of truth of *short*; (b) *medium*; (c) *long duration of outgoing calls* in various BD phases assuming that the linguistic variables are constructed individually for each patient. Figure 6. presents the resulting boxplots. As depicted in Figure 6., for example, for the attribute *short duration of outgoing calls*, the median degree of truth amounts to 0.56 for patients in euthymia and to 0.84 for patients in mania. Interestingly, the median degree of truth of the attribute *long duration of outgoing calls* amounts to 0.43 and to 0.0 for patients in euthymia and in mania, respectively.

Secondly, we construct the quantifier *most* and use it for creation of the linguistic summaries. Table 3 presents the mean degrees of truth of the personalized linguistic summaries about duration of outgoing calls in various BD states. As presented in Table 3, the mean degree of truth of the presented linguistic summaries changes depending on the state of the BD. In particular, the mean degree of truth of personalized linguistic summary *Most outgoing calls are short* for patients in euthymia amounts to 0 and for patients in mania and depression amounts to 1.0. Contrary to the approach using descriptive statistics and absolute smartphone-based measurements, the personalized linguistic summarization confirms that the smartphone-based duration of outgoing calls differs for patients in the mania state and in the euthymia. There is also difference in the mean degrees of truth between patients in depressions and euthymia and between patients in the mixed state and euthymia.

6 Conclusion and Future Work

Within this research, we developed the personalized linguistic summaries and applied them in the smartphone-based monitoring of the bipolar disorder patients. Personalized linguistic summaries are patient-dependent and disease-state-dependent. Their performance is illustrated with preliminary numerical results about duration of outgoing calls. We show that there are differences between degrees of truth of the personalized linguistic summaries in various BD phases. The personalized linguistic summariza-

tion confirmed that the smartphone-based duration of outgoing calls is different for patients in euthymia vs. patients in a disease state (mania, depression, mixed state).

It needs to be noted that this is the first application of the linguistic summarization to support the smartphone-based monitoring of the bipolar disorder. Further research is planned to verify all limitations and advantages of the proposed methodology using other smartphone-based parameters gathered within the observational study. We also plan to assess the efficiency of the personalized approach when informing about an alarming situation and to improve the formalization of concepts related to the personalized linguistic variables.

Acknowledgement

This work was partially financed from EU funds (Regional Operational Program for Mazovia) – a project entitled "Smartphone-based diagnostics of phase changes in the course of bipolar disorder" (RPMA.01.02.00-14-5706/16-00). The authors would also like to thank the researchers Anna Olwert, Karol Opara and Weronika Radziszewska for their support and advice.

References

- [1] R. Almeida, M. Lesot, B. Bouchon-Meunier, U. Kaymak, G. Moyse, Linguistic summaries of categorical time series patient data, *Proc. of FUZZ-IEEE'13* (2013) 1–8.
- [2] J. M. Alonso, A. Bugarin, E. Reiter, Natural language generation with computational intelligence, *IEEE Computational Intelligence Magazine* 12(3) (2017) 8–9.
- [3] A. Alvarez-Alvarez, G. Trivino, Linguistic description of the human gait quality, *Engineering Applications of Artificial Intelligence* 26(1) (2013) 13–23.
- [4] A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition, 5th Edition, Washington, DC, 2000, Ch. Bipolar and Related Disorders in Diagnostic.
- [5] F. E. Boran, D. Akay, R. R. Yager, An overview of methods for linguistic summarization with fuzzy sets, *Expert Systems with Applications* 61 (2016) 356–377.
- [6] R. Castillo-Ortega, N. Marín, D. Sánchez, Linguistic local change comparison of time series, *Proc. of FUZZ-IEEE'11* (2011) 2909–2915.
- [7] E. Gliddon, S. Barnes, G. Murray, E. Michalak, Online and mobile technologies for self-management in bipolar disorder: A systematic review., *Psychiatric Rehabilitation Journal* 40 (3) (2017) 309–319.
- [8] I. Grande, M. Berk, B. Birmaher, E. Vieta, Bipolar disorder, *The Lancet* 387 (10027) (2016) 1561 – 1572.
- [9] A. Grünerbl, A. Muaremi, V. Osmani, Smartphone-based recognition of states and state changes in bipolar disorder patients, *IEEE Journal of Biomedical and Health Informatics* 19(1) (2015) 140 – 148.
- [10] O. Hryniewicz, K. Kaczmarek, Bayesian analysis of time series using granular computing approach, *Applied Soft Computing* 47 (2016) 644–652.
- [11] O. Hryniewicz, K. Kaczmarek-Majer, Monitoring of time series using fuzzy weighted prediction models, in: S. Destercke, T. Denoeux, M. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), *Uncertainty Modelling in Data Science*, Vol. 832 of *Advances in Intelligent Systems and Computing*, Springer International, 2018, pp. 107–114.
- [12] J. Kacprzyk, Linguistic summarization of time series using a fuzzy quantifier driven aggregation, *Fuzzy Sets Syst* 159 (12) (2008) 1485–1499.
- [13] J. Kacprzyk, R. Yager, Linguistic summaries of data using fuzzy logic, *International Journal of General Systems* 30(2) (2001) 133–154.
- [14] J. Kacprzyk, R. Yager, J. Merigo, Towards human-centric aggregation via ordered weighted aggregation operators and linguistic data summaries: A new perspective on zadeh's inspirations, *IEEE Computational Intelligence Magazine* 14 (2019) 16–30.
- [15] J. Kacprzyk, S. Zadrozny, Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools, *Information Sciences* 173 (2005) 281–304.
- [16] K. Kaczmarek-Majer, O. Hryniewicz, Application of linguistic summarization methods in time series forecasting, *Information Sciences* 478 (2019) 580–594.
- [17] K. Kaczmarek-Majer, O. Hryniewicz, K. Opara, W. Radziszewska, J. Owsinski, S. Zadrozny, Control charts designed using model averaging approach for phase change detection in bipolar disorder, in: S. Destercke, T. Denoeux, M. Gil,

- P. Grzegorzewski, O. Hryniewicz (Eds.), *Uncertainty Modelling in Data Science*, Vol. 832 of *Advances in Intelligent Systems and Computing*, Springer International, 2018, pp. 115–123.
- [18] M. Lesot, G. Moyse, B. Bouchon-Meunier, Interpretability of fuzzy linguistic summaries, *Fuzzy Sets and Systems* 292 (2016) 307–317.
 - [19] N. Marín, D. Sanchez, On generating linguistic descriptions of time series, *Fuzzy Sets and Systems* 285 (2016) 6–30.
 - [20] J. Moreno-Garcia, J. Abián-Vicén, L. Jimenez-Linares, L. Rodriguez-Benitez, Description of multivariate time series by means of trends characterization in the fuzzy domain, *Fuzzy Sets and Systems* 285 (2016) 118–139.
 - [21] G. Moyse, M.-J. Lesot, Linguistic summaries of locally periodic time series, *Fuzzy Sets and Systems* 285 (2016) 94–117.
 - [22] V. Novak, Evaluative linguistic expressions vs. fuzzy categories, *Fuzzy Sets and Systems* 281 (2015) 73–87.
 - [23] M. D. Peláez-Aguilera, M. Espinilla, M. R. F. Olmo, J. Medina, Fuzzy linguistic protoforms to summarize heart rate streams of patients with ischemic heart disease, *Complexity* 4(1) (2019) 1–11.
 - [24] A. Ramos-Soto, A. Bugarín, S. Barro, On the role of linguistic descriptions of data in the building of natural language generation systems, *Fuzzy Sets and Systems* 285 (2016) 31–51.
 - [25] A. Wilbik, J. M. Keller, J. C. Bezdek, Linguistic prototypes for data from eldercare residents, *IEEE Transactions on Fuzzy Systems* 22 (1) (2014) 110–123.
 - [26] R. Yager, A new approach to the summarization of data, *Information Science* 28 (1) (1982) 69–86.
 - [27] L. A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, *Computers and Maths with Applications* (1983) 149–184.