This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Bipolar disorder prediction with sensor-based semi-supervised learning

D4.1 – Requirement analysis

Deliverable No.	D4.1	**Due date**	30-SEP-2022
Type	Report	**Dissemination Level**	Public
Version	1.0	**WP**	WP1
Description	Requirements mapped onto use cases defined in psychiatric scenarios		

Systems Research Institute Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

Authors

Name	Email
Katarzyna Kaczmarek-Majer	k.kaczmarek@ibspan.waw.pl
Olga Kamińska	o.kaminska@ibspan.waw.pl
Kamil Kmita	k.kmita@ibspan.waw.pl
Jakub Małecki	jakub.malecki@ibspan.waw.pl
Izabella Zadrożna	izabella.zadrozna@ibspan.waw.pl

History

Date	Version	Change
12-SEP-2022	0.1	Integrated version of the document
30-SEP-2022	1.0	Initial version of the requirement analysis

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Executive summary

This deliverable outlines the results of Task 4.1 activities, which were driven by the BIPOLAR team in cooperation with two experienced professionals in psychiatry, namely Dr Monika Dominiak from Department of Affective Disorders, Institute of Psychiatry and Neurology in Warsaw, Poland and Prof. Svetlozar Haralanov from University Hospital for Neurology and Psychiatry "St. Naum" in Sofia, Bulgaria. This interdisciplinary team collected requirements for BIPOLAR from literature and in consultancy with the practitioners in psychiatry. Next, requirements were mapped onto use cases defined in psychiatric scenarios prepared in Task 1.1. Five psychiatric scenarios need to be addressed by the BIPOLAR package as described in Deliverable 1.2.

The analysis of each psychiatric scenario allowed for the definition and refinement of the project's scope. It also led to better understanding and agreement between the team members about the requirements that need to be addressed by the BIPOLAR project. Based on the common understanding of the functionality of the BIPOLAR package under development, its functionality was documented in the form of use cases, which describe the interactions between actors (persons, devices or digital entities), the assumptions and the expected outcomes after the execution of a flow of actions. Further functional and pilot specific requirements were identified based on the expertise and experience of all team members. This report includes requirements for 5 psychiatric scenarios grouped into 2 pilots of BIPOLAR, that are Acoustic Data Pilot (ADP): Acoustic data collected from smartphones of bipolar disorder patients and Locomotor Data Pilot (LDP): Locomotor data collected from sensor of bipolar disorder and unipolar depression patients.

Requirements will be divided with respect to selected categories, namely functional and pilot-specific requirements. In the future, BIPOLAR's key performance indicators will be specified based on the identified use cases and requirements in order to be able to measure the outcome of the pilots and provide input for the technical and use-case evaluation of the project results.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Table of Contents

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

List of acronyms

Acronym	Explanation
BIPOLAR	Bipolar disorder prediction with sensor-based semi-supervised learning project
BD	Bipolar disorder
ADP	Acoustic Data Pilot
LDP	Locomotor Data Pilot
SSFCM	Semi-supervised fuzzy c-means algorithm

1.Introduction

1.1. About this document

This deliverable outlines the requirements gathered as part of Tasks 4.1 activities which were driven by the BIPOLAR team in cooperation with advisors representing the medical domain. BIPOLAR project delivers an open software package containing a set of computational methods that support engineering of information from sensors and semi-supervised learning to provide accurate prediction of shifts from euthymia to depression, mania and mixed state. Requirements for this software package will be divided with respect to the following categories:

- Functional requirements;
- Pilot-specific requirements;

These requirements are closely related to the psychiatric scenarios delivered as Deliverable 1.1 and will be listed in the remaining of this document.

Various definitions exist of what a requirement is. The following two definitions of ISO and INCOSE are adapted for this report:

- "A requirement is Statement that identifies a product (includes product, service, or enterprise) or process operational, functional, or design characteristic or constraint, which is unambiguous, testable or measurable, and necessary for product or process acceptability." (ISO/IEC 2007)[1]
- "A requirement is a statement that identifies a system, product or process characteristic or constraint, which is unambiguous, clear, unique, consistent, stand-alone (not grouped), and verifiable, and is deemed necessary for stakeholder acceptability." (INCOSE 2010)[2]

The Volere methodology[3] was applied to define the use cases and requirements under various scenarios. In BIPOLAR project, all scenarios are dedicated to support a particular medical domain that is psychiatry,

[1] ISO/IEC. 2007. Systems and Software Engineering -- Recommended Practice for Architectural Description of SoftwareIntensive Systems. Geneva, Switzerland: International Organization for Standards (ISO)/International Electrotechnical Commission (IEC), ISO/IEC 42010:2007.

[2] INCOSE. 2010. Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities. Version 3.2.1. San Diego, CA, USA: International Council on Systems Engineering (INCOSE), INCOSE-TP2003-002-03.2.1: 362

[3] Volere Requirements: How to Get Started
https://www.volere.org/wpcontent/uploads/2018/12/VolereGettingStarted.pdf

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

and thus, they are called psychiatric scenarios in the remaining of this document. Psychiatric scenario identified in the BIPOLAR project were described using the template according to as defined in Deliverable 1.1. Use case specification includes (among other): (i) actors (persons, devices or digital entities) that interact and participate in the use cases; (ii) requirements, assumptions and/or pre-conditions to be satisfied for the use case to be performed; (iii) flow of events between actors, and sequences of interactions focusing on differences from current operations, and (iv) expected outcomes after the use case execution.

2. Functional requirements

Functional requirements are the fundamental subject matter of the system and are measured by concrete means like: data values, decision-making logic, and algorithms. The main functionality requirements for BIPOLAR package are defined in Table 1. All of them have highest priority (must be delivered within this project).

Table 1. Functional requirements for the BIPOLAR package

ID	Name	Description
FR1	Affective state recognition	BIPOLAR package **must provide at least three selected top performing supervised learning methods i**mplemented as benchmarks for affective state recognition, such as e.g., RF, XGBoost, SVM, etc. Affective state recognition represents a binary or a multiple class classification problem.
FR2	Label extrapolation	BIPOLAR package **must provide** a function for extrapolating in time labels assigned by psychiatrist to sensor data. This extrapolation must be easy to be configured by the users of the package.
FR3	Calculating confidence of labels	BIPOLAR package **must provide** a method for calculating confidence of observations. The idea is to intelligently combine unsupervised data from sensors with labels from external label source and to estimate uncertainty related with this process. The uncertainty is expressed by a confidence factor assigned to each observation that became supervised in the result of the aforementioned process. Thus, the confidence factor is a fraction ranging from 0 to 1 which describes a level of confidence we assign to the supervised assumption that a given observation belongs to the certain class (e.g., that the specific call belongs to the "depression" phase of bipolar disorder).
FR4	Feature selection	BIPOLAR package **must provide** a uncertainty-aware method for feature selection.
FR5	Semi-supervised affective state prediction	BIPOLAR package **must provide** a method for semi-supervised prediction of the affective states. This could be either binary class problem or a multiclass problem.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

FR6	Generation of dashboard about patient's status	BIPOLAR package **must provide** a method for dashboard generation. Dashboard is aimed to present in an interactive way about (1) basic patients' information gathered during the visit and the (2) data statistics.
FR7	Generation of dashboard about models'performance	BIPOLAR package **must provide** a method for dashboard generation. Dashboard is aimed to present in an interactive way about results received from executed semi-supervised algorithms with their visualizations.

BIPOLAR will deliver at least one algorithm for each task described in Table 1. Algorithms have been chosen based on the task requirements of the two pilots described in Deliverable 1.1 dedicated to psychiatric scenarios as well as basic functionality described in WP1, WP2, WP3 and WP4 of the project.

All functions will be documented according to programming in R language best practices and standards. Additionally, Trello work management tool will be used to save results of online discussions and efficiently track workflows and tasks assigned to the team. With Trello teams can ideate plans, collaborate on projects, organize workflows, and track progress in a visual and productive way. Trello manages the milestones and the day-to-day tasks.

We now discuss requirements specific for the identified psychiatric scenarios in pilots.

2. Pilot-specific requirements

Table 2 gathers requirements considered for demonstrating the functionalities of BIPOLAR in Pilot 1 – Acoustic Data Pilot (ADP) and Pilot 2 –Locomotor Data Pilot (LDP).

Table 2. Pilot specific requirements for the BIPOLAR package

ID	Name	Description
ADPR1	Extended label extrapolation	BIPOLAR package **must provide** a method for annotating unsupervised data from one data source (e.g., data about voice characteristics from a phone call collected by a smartphone application) with information from a different data source (e.g. a psychiatrist's assessment about the current phase of the bipolar disorder) based on common factors (e.g. patient's identity and the time domain). In the context of psychiatric disease monitoring, the merging of the data is performed based on the "ground truth period" – a time frame around the visit considered as viable by the expert.
ADPR2	Estimation of label uncertainty based on predefined distribution	BIPOLAR package **must provide** a method for assignment of uncertainty based on some predefined distribution, e.g., Gaussian. Uncertainty is expressed by a confidence factor derived for each observation that

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

		became supervised in a process of label extrapolation. The confidence factor, a fraction ranging from 0 to 1, is estimated using the datasets provided by the user. The methodology of estimation is based on SSFCM algorithm as described in paper of [Kmita et al., 2022].[4]
ADPR3	Acoustic feature aggregation	BIPOLAR package **must provide** a method for automated calculation of aggregates, e.g., to aggregate individual frames to the level of recordings (calls).
LDPR1	Semi-supervised prediction of affected individuals vs. healthy controls	BIPOLAR package **must provide** a method for semi-supervised prediction of healthy controls vs. affected patients based on the locomotor data. This would be a binary classification problem.

[4] K. Kmita, G. Casalino, G. Castellano, O. Hryniewicz and K. Kaczmarek-Majer, "Confidence path regularization for handling label uncertainty in semi-supervised learning: use case in bipolar disorder monitoring," 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2022, pp. 1-8, doi: 10.1109/FUZZ-IEEE55066.2022.9882759.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.