Elaine Fehrman · Vincent Egan ·
Alexander N. Gorban · Jeremy Levesley ·
Evgeny M. Mirkes · Awaz K. Muhammad

Personality Traits and Drug Consumption

A Story Told by Data

Springer

Personality Traits and Drug Consumption

Elaine Fehrman · Vincent Egan ·
Alexander N. Gorban · Jeremy Levesley ·
Evgeny M. Mirkes · Awaz K. Muhammad

Personality Traits and Drug Consumption

A Story Told by Data

Springer

Elaine Fehrman
Men's Personality Disorder and National Women's Directorate
Rampton Hospital
Retford, UK

Vincent Egan
Department of Psychiatry and Applied Psychology
University of Nottingham
Nottingham, UK

Alexander N. Gorban
Department of Mathematics
University of Leicester
Leicester, UK

Jeremy Levesley
Department of Mathematics
University of Leicester
Leicester, UK

Evgeny M. Mirkes
Department of Mathematics
University of Leicester
Leicester, UK

Awaz K. Muhammad
Department of Mathematics
University of Leicester
Leicester, UK

Department of Mathematics
College of Education
Salahaddin University-Erbil
Kurdistan Region, Iraq

ISBN 978-3-030-10441-2 ISBN 978-3-030-10442-9 (eBook)
https://doi.org/10.1007/978-3-030-10442-9

Library of Congress Control Number: 2018965426

Mathematics Subject Classification (2010): 62-07, 62P15, 62H30, 62H25

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
Introduction

Each set of data can tell us a story. Our mission is to extract this story from the data and translate it into more readily accessible human language. There are a number of tools for such a translation. To prepare this story, we have to collect data, to ask interesting questions, and to apply all the possible data mining technical tools to find the answers. Then, we should verify the answers, exclude spurious (overoptimistic) correlations and patterns, and tell the story to users.

The topic of mining interesting knowledge remains very intriguing. Many researchers have approached this problem from a plethora of different angles. One of the main ideas in these approaches has been information gain (the more information gain there is, the more interesting the result is). Nevertheless, we need a good understanding of what makes patterns that are found interesting from the end-user's point of view. Here, various perspectives might be involved, from practical importance to aesthetic beauty. The extraction of deep and interesting knowledge from data was formulated as an important problem for the 5th IEEE International Conference on Data Mining (ICDM 2005) [1]. Nowadays, the fast growth of the fields of data science and machine learning provides us with many tools for answering such questions, but the art of asking interesting questions still requires human expertise.

The practical importance of the problem of evaluating an individual's risk of consuming and/or abusing drugs cannot be underestimated [2]. One might well ask how this risk depends on a multitude of possible factors [3]? The linking of personality traits to risk of substance use disorder is an enduring problem [4]. Researchers return again and again to this problem following the collection of new data, and with new questions.

How do personality, gender, education, nationality, age, and other attributes affect this risk? Is this dependence different for different drugs? For example, does the risk of ecstasy consumption and the risk of heroin consumption differ for different personality profiles? Which personality traits are the most important for evaluation of the risk of consumption of a particular drug, and are these traits different for different drugs? These questions are the focus of our research.

The data set we collected contains information on the consumption of 18 central nervous system psychoactive drugs, by 2,051 respondents (after cleaning, 1,885

E. Fehrman et al., *Personality Traits and Drug Consumption*,
https://doi.org/10.1007/978-3-030-10442-9_1

participants remained, male/female = 943/942). The database is available online [5, 6].

The questions we pose above have been reformulated as classification problems, and many well-known data mining methods have been employed to address these problems: decision trees, random forests, k-nearest neighbours, linear discriminant analysis, Gaussian mixtures, probability density function estimation using radial basis functions, logistic regression, and naïve Bayes. For data preprocessing, transformation, and ranking, we have used methods such as polychoric correlation, nonlinear Categorical Principal Component Analysis (CatPCA), sparse PCA, and original double Kaiser's feature selection.

The main results of the work are:

- The presentation and descriptive analysis of a database with information on 1,885 respondents and their usage of 18 drugs.
- Demonstration that the personality traits (Five-Factor Model [7], impulsivity, and sensation-seeking) together with simple demographic data give the possibility of predicting the risk of consumption of individual drugs with sensitivity and specificity above 70% for most drugs.
- The construction of the best classifiers and most significant predictors for each individual drug in question.
- Revelation of significantly distinct personality profiles for users of different drugs; in particular, groups of heroin and ecstasy users are significantly different in Neuroticism (higher for heroin), Extraversion (higher for ecstasy), Agreeableness (higher for ecstasy), and Impulsivity (higher for heroin); groups of heroin and benzodiazepine users are significantly different in Agreeableness (higher for benzodiazepines), Impulsivity (higher for heroin), and Sensation-Seeking (higher for heroin); groups of ecstasy and benzodiazepine users are significantly different in Neuroticism (higher for benzodiazepines), Extraversion (higher for ecstasy), Openness to Experience (higher for ecstasy), and Sensation-Seeking (higher for ecstasy).
- The discovery of three correlation pleiades of drugs; these are the clusters of drugs with correlated consumption centred around heroin, ecstasy, and benzodiazepines. The correlation pleiades should include the mini-sequences of drug involvement found in longitudinal studies [8] and aim to serve as maps for analysis of different patterns of influence.
- The development of risk map technology for the visualisation of the probability of drug consumption.

Four of the authors (ANG, JL, EMM, and AKM) are applied mathematicians, and two are psychologists (EF and VE). Data were collected by EF and processed by EMM and AKM. The psychological framework for this study was developed by EF and VE, and the analytic methodology was selected and developed by ANG and EMM. The final results were critically analysed and described by ANG, JL, EMM, and AKM from the data mining perspective, and EF and VE provided the psychological interpretation and conceptualisation.

For psychologists, the book gives a new understanding of the relationship between personality traits and the usage of 18 psychoactive substances, provides a new openly available database for further study, and presents many useful methods of data analysis. For applied mathematicians and statisticians, the book details a case study in a fascinating area of application, exemplifying the use of various data mining methods in such scenarios.

This book is aimed at advanced undergraduates or first-year Ph.D. students, as well as researchers and practitioners in data analysis, applied mathematics, and psychology. No previous knowledge of machine learning, advanced data mining concepts, or psychology of personality is assumed. Familiarity with basic statistics and some experience of the use of probability is helpful, as well as some basic understanding of psychology. Two books [9, 10] include all the necessary prerequisites (and much more). Linear discriminant analysis (LDA), principal component analysis (PCA), and decision trees (DT) are systematically employed in the book. Therefore, it may be useful to refresh the knowledge of these classical methods using the textbook [10], which is concentrated more on the applications of the methods and less on the mathematical details.

A preliminary report of our work was published as an arXiv e-print in 2015 [11] and presented at the Conference of International Federation of Classification Societies 2015 (IFCS 2015) [12].

This book is not the end of the story told by the data. We will continue our work and try to extract more interesting knowledge and patterns from the data. Moreover, we are happy for you, the readers, to join us in this adventure. We believe that every large annotated data set is a treasure trove and that there is an abundance of interesting knowledge to discover from them. We have published our database online [5, 6] and invite everybody to use it for their own projects, from B.Sc. and M.Sc. level to Ph.D., or just for curiosity-driven research. We would be very happy to see the fascinating outcomes of these projects.

References

1. Yang, Q., Wu, X.: 10 challenging problems in data mining research. Int. J. Inf. Technol. Decis. Mak. **5**(04), 597–604 (2006). https://doi.org/10.1142/S0219622006002258
2. United Nations Office on Drugs and Crime: World Drug Report 2016. United Nations, New York. http://www.unodc.org/wdr2016/ (2016). Accessed 27 Dec 2017
3. Hawkins, J.D., Catalano, R.F., Miller, J.Y.: Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention. Psychol. Bull. **112**(1), 64–105 (1992). https://doi.org/10.1037/0033-2909.112.1.64
4. Kotov, R., Gamez, W., Schmidt, F., Watson, D.: Linking big personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. Psychol. Bull. **136**(5), 768–821 (2010). https://doi.org/10.1037/a0020327
5. Fehrman, E., Egan, V.: Drug consumption, collected online March 2011 to March 2012, English-speaking countries. ICPSR36536-v1. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI (2016). Deposited by Mirkes, E.M. https://doi.org/10.3886/ICPSR36536.v1

6. Fehrman, E., Egan, V., Mirkes, E.M.: Drug consumption (quantified) data set. UCI Machine Learning Repository (2016). https://archive.ics.uci.edu/ml/datasets/Drug+consumption+
7. Costa, P.T., MacCrae, R.R.: Revised NEO-Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual. Psychological Assessment Resources, Odessa, FL (1992)
8. Newcomb, M.D., Bentler, P.M.: Frequency and sequence of drug use: a longitudinal study from early adolescence to young adulthood. J. Drug Educ. **16**(2), 101–120 (1986). https://doi.org/10.2190/1VKR-Y265-182W-EVWT
9. Corr, P.J., Matthews, G. (eds.): The Cambridge Handbook of Personality Psychology. Cambridge University Press, New York (2009). https://doi.org/10.1017/cbo9780511596544
10. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning. Springer Texts in Statistics, vol. 103. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7138-7
11. Fehrman, E., Muhammad, A.K., Mirkes, E.M., Egan, V., Gorban, A.N.: The five factor model of personality and evaluation of drug consumption risk. https://arxiv.org/abs/1506.06297 (2015). Accessed 27 Dec 2017
12. Fehrman, E., Muhammad, A.K., Mirkes, E.M., Egan, V., Gorban, A.N.: The five factor model of personality and evaluation of drug consumption risk. In: Palumbo, F., Montanari, A., Vichi, M. (eds.) Data Science, Studies in Classification, Data Analysis, and Knowledge Organization, pp. 215–226. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-55723-6_18

Chapter 2
Drug Use and Personality Profiles

Abstract Drug use disorder is characterised by several terms: addiction, dependence, and abuse. We discuss the notion of psychoactive substance and relations between the existing definitions. The personality traits which may be important for predisposition to use of drugs are introduced: the Five-Factor Model, impulsivity, and sensation-seeking. A number of studies have illustrated that personality traits are associated with drug consumption. The previous pertinent results are reviewed. A database with information on 1,885 respondents and their usage of 18 drugs is introduced. The results of our study are briefly outlined: the personality traits (Five-Factor Model, impulsivity, and sensation-seeking) together with simple demographic data make possible the prediction of the risk of consumption of individual drugs; personality profiles for users of different drugs. In particular, groups of heroin and ecstasy users are significantly different; there exist three correlation pleiades of drugs. These are clusters of drugs with correlated consumption, centred around heroin, ecstasy, and benzodiazepines.

Keywords Psychoactive drugs · Drug users · Personality traits · Five-Factor Model · Data description

2.1 Definitions of Drugs and Drug Usage

Since Sir Karl Popper, it has become a commonplace opinion in the philosophy of science that the 'value' of definitions, other than for mathematics, is generally unhelpful. Nevertheless, for many more practical spheres of activity, from jurisprudence to health planning, definitions are necessary to impose theoretical boundaries on a subject, in spite of their incompleteness and their tendency to change with time. This applies strongly to definitions of drugs and drug use.

Following the standard definitions [1],

- A *drug* is a 'chemical that influences biological function (other than by providing nutrition or hydration)'.
- A *psychoactive drug* is a 'drug whose influence is in a part on mental functions'.

E. Fehrman et al., *Personality Traits and Drug Consumption*,
https://doi.org/10.1007/978-3-030-10442-9_2

- An *abusable psychoactive drug* is a ‘drug whose mental effects are sufficiently pleasant or interesting or helpful that some people choose to take it for a reason other than to relieve a specific malady’.

In our study, we use the term ‘drug’ for abusable psychoactive drug regardless of whether it is illicit or not. While legal substances such as sugar, alcohol, and tobacco are probably responsible for far more premature death than illegal recreational drugs [2], the social and personal consequences of recreational drug use can be highly problematic [3].

Use of drugs introduces risk into a life across a broad spectrum; it constitutes an important factor for increasing risk of poor health, along with earlier mortality and morbidity, and has significant consequences for society [4, 5]. Drug consumption and addiction constitutes a serious problem globally. Though drug use is argued by civil libertarians to be a matter of individual choice, the effects on an individual of drug use, such as greater mortality or lowered individual functioning, suggest that drug use has social and interpersonal effects on individuals who have not chosen to use drugs themselves.

Several terms are used to characterise drug use disorder: addiction, dependence, and abuse. For a long time, ‘substance abuse’ and ‘substance dependence’ were considered as separate disorders. In 2013, The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) joined these two diagnoses into ‘Substance Use Disorder’ [6]. This is a more inclusive term used to identify persons with substance-related problems. More recently, abuse and dependence have been defined on a scale that measures the time and degree of substance use. Criteria are provided for substance use disorder, supplemented by criteria for intoxication, withdrawal, substance-/medication-induced disorders, and unspecified substance-induced disorders, where relevant. Abuse can be considered as the early stage of substance use disorder.

In our study, we differentiate the substance users on the basis of recency of use but do not identify existence and depth of the substance dependence.

For prevention and effective care of substance use disorder, we need to identify the risk factors and develop methods for their evaluation and control [7].

2.2 Personality Traits

Sir Francis Galton (1884) proposed that a lexical approach in which one used dictionary definitions of dispositions could be a means of constructing a description of individual differences (see [8]). He selected the personality-descriptive terms and stated the problem of their interrelations. In 1934, Thurstone [9] selected 60 adjectives that are in common use for describing people and asked each of 1300 respondents to think of a person they knew well and to select the adjectives that can best describe this person. After studying the correlation matrix, he found that *five* factors are sufficient to describe this choice.

There have been many versions of the Five-Factor Model proposed since Thurston [10], for example:

- Surgency, agreeableness, dependability, emotional stability, and culture;
- Surgency, agreeableness, conscientiousness, emotional stability, and culture;
- Assertiveness, likeability, emotionality, intelligence, and responsibility;
- Social adaptability, conformity, will to achieve, emotional control, and inquiring intellect;
- Assertiveness, likeability, task interest, emotionality, and intelligence;
- Extraversion, friendly compliance, will to achieve, neuroticism, and intellect;
- Power, love, work, affect, and intellect;
- Interpersonal involvement, level of socialisation, self-control, emotional stability, and independence.

There are also systems with different numbers of factors (three, seven, etc.). The most important three-factor system is Eysenck's model comprising extraversion, psychoticism, and neuroticism.

Nowadays, after many years of research and development, psychologists have largely agreed that the personality traits of the modern Five-Factor Model (FFM) constitutes the most comprehensive and adaptable system for understanding human individual differences [11]. The FFM comprises Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C).

The five traits can be summarised as follows:

N *Neuroticism* is a long-term tendency to experience negative emotions such as nervousness, tension, anxiety, and depression (associated adjectives [12]: anxious, self-pitying, tense, touchy, unstable, and worrying).

E *Extraversion* is manifested in characters who are outgoing, warm, active, assertive, talkative, and cheerful; these persons are often in search of stimulation (associated adjectives: active, assertive, energetic, enthusiastic, outgoing, and talkative).

O *Openness to experience* is associated with a general appreciation for art, unusual ideas, and imaginative, creative, unconventional, and wide interests (associated adjectives: artistic, curious, imaginative, insightful, original, and wide interest).

A *Agreeableness* is a dimension of interpersonal relations, characterised by altruism, trust, modesty, kindness, compassion, and cooperativeness (associated adjectives: appreciative, forgiving, generous, kind, sympathetic, and trusting).

C *Conscientiousness* is a tendency to be organised and dependable, strong-willed, persistent, reliable, and efficient (associated adjectives: efficient, organised, reliable, responsible, and thorough).

Individuals low on the A and C trait dimensions have less incidence of the reported attributes, so, for example, lower Agreeableness is associated with greater antisocial behaviour [13].

2.3 How Many Inputs Do the Predictive Models Have: 5, 30, 60, or 240?

The NEO PI-R questionnaire was specifically designed to measure the FFM of personality [11]. It provides scores corresponding to N, E, O, A, and C (‘domain scores’). The NEO PI-R consists of 240 self-report items answered on a five-point scale, with separate scales for each of the five domains. Each scale consists of six correlated subscales (‘facets’). A list of the facets within each domain is presented in the first column of Table 2.1.

There are several versions of the FFM questionnaire: NEO PI-R with 240 questions (‘items’), 30 facets, and five domains; the older NEO-FFI with 180 items, etc. A shorter version of the Revised NEO Personality Inventory (NEO PI-R), the NEO Five-Factor Inventory (NEO-FFI), has 60 items (12 per domain and no facet structure) selected from the original items [11]. This shorter questionnaire was revised [14] after Egan et al. demonstrated that the robustness of the original version should be improved [15]. NEO-FFI was designed as a brief instrument that provides estimates of the factors for use in exploratory research.

The values of the five factors are used as inputs in numerous statistical models for prediction, diagnosis, and risk evaluation. These models are employed in psychology, psychiatry, medicine, education, sociology, and many other areas where personality may be important. For example [16], academic performance at primary school was found to significantly correlate with Emotional Stability (+), Agreeableness (+), Conscientiousness (+), and Openness to Experience (+) (the sign of correlations is presented in parentheses). Success in primary school is also significantly and highly correlated with intelligence (+), the Pearson correlation coefficient $r > 0.5$. For higher academic levels, correlations of Academic Performance with Emotional Stability, Agreeableness, and Openness significantly decreases ($r \lessapprox 0.1$). Correlation with Intelligence also decreases by two or more, but correlation with Conscientiousness remains almost the same for all academic levels ($r \approx 0.21$–0.28). Correlations between Conscientiousness and Academic Performance were largely independent of Intelligence. This knowledge can be useful for educational professionals and parents.

Another example demonstrates how personality affects career success [17]. Extraversion was related positively to salary level, promotions, and career satisfaction, and Neuroticism was related negatively to career satisfaction. Agreeableness was related negatively only to career satisfaction, and Openness was related negatively to salary level. There was a significant negative relationship between Agreeableness and salary for individuals in people-oriented occupations (with direct interaction with clients, for example), but no such relationships were found in occupations without a strong ‘people’ component. At the same time, Agreeableness is positively correlated with performance in jobs involving teamwork (interaction with co-workers) [18]. These results are of interest to human resources departments.

Most of the statistical models use the values of five factors (N, E, O, A, and C) as the inputs and produce assessment, diagnosis, recommendations, or prognosis as the outputs (Fig. 2.1a). For the NEO PI-R questionnaire, this means that we take the

Table 2.1 FFM facet trait predictor set for DSM-IV PD [19, 22]

FFM	PAR	SZD	SZT	ATS	BDL	HST	NAR	AVD	DEP	OBC
Neuroticism										
Anxiety			$\Uparrow$		$\Uparrow$			$\Uparrow$	$\Uparrow$	
Angry hostility	$\Uparrow$			$\Uparrow$	$\Uparrow$		$\Uparrow$			
Depression					$\Uparrow$	$\Uparrow$		$\Uparrow$		
Self-consciousness			$\Uparrow$			$\Uparrow$	$\Uparrow$	$\Uparrow$	$\Uparrow$	
Impulsiveness					$\Uparrow$					
Vulnerability					$\Uparrow$			$\Uparrow$	$\Uparrow$	
Extraversion										
Warmth		$\Downarrow$	$\Downarrow$			$\Uparrow$			$\Uparrow$	
Gregariousness		$\Downarrow$	$\Downarrow$			$\Uparrow$		$\Downarrow$		
Assertiveness								$\Downarrow$	$\Downarrow$	$\Uparrow$
Activity										
Excitement seeking				$\Uparrow$		$\Uparrow$		$\Downarrow$		
Positive emotions		$\Downarrow$	$\Downarrow$			$\Uparrow$				
Openness to Experience										
Fantasy			$\Uparrow$			$\Uparrow$	$\Uparrow$			
Aesthetics										
Feelings		$\Downarrow$				$\Uparrow$				
Actions			$\Uparrow$							
Ideas			$\Uparrow$							
Values										$\Downarrow$

(continued)

Table 2.1 (continued)

FFM	PAR	SZD	SZT	ATS	BDL	HST	NAR	AVD	DEP	OBC
Agreeableness										
Trust	⇓		⇓		⇓	⇑			⇑	
Straightforwardness	⇓			⇓						
Altruism				⇓			⇓		⇑	
Compliance	⇓			⇓	⇓				⇑	⇓
Modesty							⇓		⇑	
Tender-mindedness				⇓			⇓			
Conscientiousness										
Competence					⇓					⇑
Order										⇑
Dutifulness				⇓						⇑
Achievement striving							⇑			⇑
Self-discipline				⇓						
Deliberation				⇓						

⇑ = High values; ⇓ = Low values; personality disorders: PAR = Paranoid; SZD = Schizoid; SZT = Schizotypal; ATS = Antisocial; BDL = Borderline; HST = Histrionic; NAR = Narcissis-Narcissistic; AVD = Avoidant; DEP = Dependent; OBC = Obsessive–compulsive

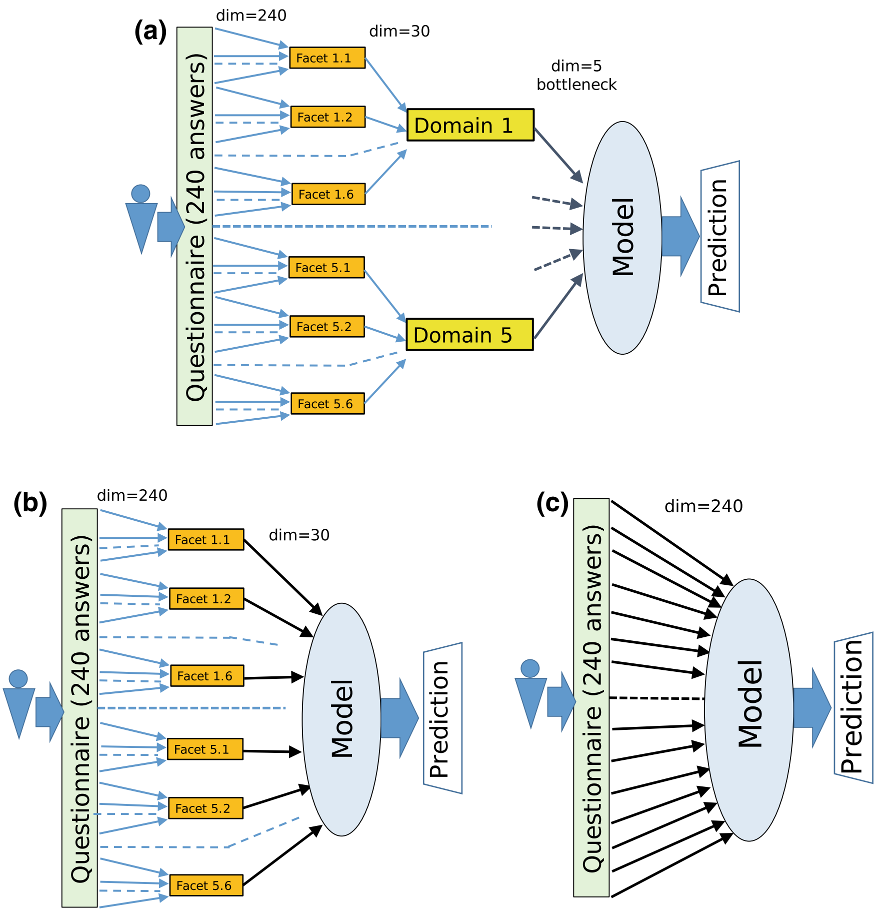

Fig. 2.1 Three types of predictive models based on the FFM NEO PI-R questionnaire: **a** statistical models which uses five FFM inputs prepared by the standard FFM procedure, **b** facet-based predictors [19, 22], and **c** direct predictors, which avoid the step of explicit diagnosis and work with multidimensional raw input information (usually, artificial intelligence models like neural networks [20, 21])

240 inputs, transform them into 30 facet values, then transform these 30 numbers into five factors, and use these five numbers as the inputs for the statistical or, more broadly, data analytic model. To construct this model with five inputs and the desired outputs, one should use data with known answers and supervising learning (or, more narrowly, various regression and classification models). The crucial question arises: is it true that for all specific diagnosis, assessment, prognosis, and recommendation problems the facets should be linearly combined with the same coefficients?

An alternative version is the facet trait model (Fig. 2.1b), where combining of facets into the final output depends on the problem and data [19]. We can go further and consider flexible combinations of the raw information, the questionnaire answers for each problem, and data set (Fig. 2.1c) [20, 21].

One of the most developed areas of FFM application is psychiatry and psychology, for example, for the assessment of personality psychopathology. The facet trait model created for 10 personality disorders [19, 22] demonstrates that optimal combinations of facets into predictors are not uniform inside the domains (Table 2.1). Some facets are more important for assessment than the others, and the selection of important facets depends on the specific personality disorder (see Table 2.1). Nevertheless, there are almost no internal contradictions inside domains in Table 2.1: for almost each domain and any given disorder, all significant facets have the same sign of deviation from the norm: either all have higher values ($\Uparrow$), or all have lower values ($\Downarrow$). The only exclusion is the contradiction between facets 'Warmth' and 'Assertiveness' from the domain 'Extraversion': both are important for the diagnosis 'Dependent', but for this diagnosis 'Warmth' is expected to be higher than average and 'Assertiveness' is expected to be lower.

In 1995, Dorrer and Gorban with co-authors [20] employed neural network technology and the original software library MultiNeuron [21] for direct prediction of human relations on the basis of raw questionnaire information. A specially reduced personality questionnaire with 91 questions was prepared. The possible answers to each question were: 'yes', 'do not know', and 'no', which were coded as $+1$, 0, and -1, correspondingly. The neural networks (committees of six networks of different architecture) were prepared to predict results of sociometry of relations between university students inside an academic group. Neural networks had to predict students' answers to the sociometric question: 'To what degree would you like to work in your future profession with this group member?' The answer was supposed to be given as a 10-point estimate (0—most negative attitude to a person as a would-be co-worker, 10—maximum positive). The status and expansivity of each group member were evaluated from the answers to these questions. Sociometric status is a measurement that reflects the degree to which someone is liked or disliked by their peers from a group. Social expansivity is the tendency of a group members to choose and highly evaluate many others. These two characteristics were used as elements of neural network output vector for each person. The inputs were 91 answers of this person to the personality questionnaire. The neural networks were trained on data from several academic groups and tested on academic groups never seen before.

The 91 questions from the questionnaire were ranked by importance for the neural networks prediction. Cross-validation showed that reduction of the questionnaire to 46 questions (the empirically optimal number in these experiments) gave the best prediction result. Committees of networks always gave better results than a single network.

While such (relatively novel) systems are often more accurate, they are more costly in two ways: they are hungrier in terms of data requirements and computational resources.

In this book, we focus on the classical systems with explicitly measured personality, which have a bottleneck of five (or seven) factors (Fig. 2.1a). Nevertheless, modern development of artificial intelligence and neural network systems ensures