Notes on Intuitionistic Fuzzy Sets
Print ISSN 1310–4926, Online ISSN 2367–8283
2022, Volume 28, Number 4, Pages 464–474
DOI: 10.7546/nifs.2022.28.4.464-474

InterCriteria Analysis as a tool for analyzing Big Data datasets: Case study of 2021 national statistics of Bulgarian system of higher education

Veselina Bureva [1,*], Petar R. Petrov [2], Vassia Atanassova [3], Ivo Umlenski [4]

[1] "Prof. Asen Zlatarov" University
1 "Prof. Yakimov" Blvd., Burgas – 8010, Bulgaria
e-mail: vbureva@btu.bg

[2] "Prof. Asen Zlatarov" University
1 "Prof. Yakimov" Blvd., Burgas – 8010, Bulgaria
e-mail: peshopbs2@gmail.com

[3] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences
105 "Acad. Georgi Bonchev" Str, Sofia – 1113, Bulgaria
e-mail: vassia.atanassova@biomed.bas.bg

[4] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences
105 "Acad. Georgi Bonchev" Str, Sofia – 1113, Bulgaria
e-mail: ivo@biomed.bas.bg

Received: 10 October 2022 **Revised:** 12 November 2022 **Accepted:** 1 December 2022

Abstract: In the paper, the intuitionistic fuzzy sets based InterCriteria Analysis (ICA) and big data systems are applied to the 2021 dataset of the National Statistical Institute with respect to the Bulgarian system of higher education. The results are produced using the Big Data platform *Hortonworks* and the ICA software *ICrAData*. Discussion is presented on the discovered relationships regarding tendencies between enrolled and graduated students, and teaching staff.
Keywords: InterCriteria Analysis, Big Data, Students, Education.
2020 Mathematics Subject Classification: 03E72.

* Paper presented at the International Workshop on Intuitionistic Fuzzy Sets, founded by Prof. Beloslav Riečan, 2 December 2022, Banská Bystrica, Slovakia.

1 Introduction

Big data analytics are implemented in many different big data platforms. Desktop or web-based applications can be used to analyze big data datasets using connection to the big data files or databases (warehouses/data lakes/data lakehouses). It is known that one of the main characteristics of big data is the variety. The data sources can contain structured data, semi-structured data and unstructured data. The information can be stored in the form of relational database, NoSQL database, text files, CSV files, LOG files, XML, etc., [9]. Thereafter these datasets can be analyzed using different tools [2, 8, 9]. Big data has been applied to problems in the field of education and e-learning [22, 24–26].

InterCriteria Analysis (ICA) is a method allowing making decisions on the basis of discovery of correlations between data using the apparatus of intuitionistic fuzzy sets [1, 3, 4, 30] (online repository of publications on ICA is given at [17]). In ICA, data arrays with measurements of multiple objects against multiple criteria are processed to calculate for every couple of criteria the intuitionistic fuzzy pairs that represent the levels of positive consonance, negative consonance and dissonance between that couple. In the case of considerably high (with respect to certain threshold values) positive consonance between a couple of criteria, ICA can enable reduction of criteria, specifically those that are slower, more subjective, more expensive, or in any other way more unfavourable for measurement, which results in delaying or raising the cost of the overall process of decision making.

Computations necessary to produce the ICA analysis of data follow a simple algorithm [3], that has several freely available implementations so far [18]: ICrAData software (in Java, C++) developed by N. Ikonomov [16], ICDM software (in C++) developed by D. Mavrov [20, 21], and a Matlab script developed by P. Vassilev and O. Roeva. These applications process XLS, XLSX and tab-delimited CSV files, and ICrAData further allows manual input of data. The basic algorithm for computation of ICA results has been an object of investigation [27], and further various methods of establishing the threshold values in ICA have been elaborated [5, 6, 7, 15].

InterCriteria Analysis has been successfully applied to decision making problems and datasets of a wide range of areas like medicine and healthcare, biotechnologies, ecology, economics and finance, industry, sports, artificial intelligence [13, 17]. Specifically, ICA has been applied to analyzing various aspects of the system of higher education [19], the Bulgarian education system [11, 12] and the education systems in other countries [10, 14, 23, 28. 29].

2 InterCriteria Analysis as a tool for analyzing Big Data datasets

In the current investigation we analyze real data from the website of the National Statistical Institute of Republic of Bulgaria for “Education and Lifelong Learning”. The dataset for 2021 year is loaded in the Hortonworks Big Data Platform in the form of files of HDFS (Hadoop File System). Apache Ambari is a tool for managing the Apache Hadoop clusters in Hortonworks data platform (HDP). Obviously, the first three services, visible in Ambari are HDFS, YARN

and MapReduce2. The applications Tez, Hive (data warehouse), Hbase (database), Pig, Sqoop, Oozie, Zookeper, Storm, Infra Solr, Atlas, Kafka (streams), Knox, Ranger, Spark2, Zeppelin Notebook, Data Analytics Studio, Druid (OLAP) and Superset are also provided in HDP (Figure 1). The services can be stopped or started depending of the desired configuration.

The login on the Hortonworks big data platform is made using *maria_dev* user. The folder "Students Files" is created. Files for education and lifelong learning are uploaded into the HDFS (Hadoop File System) (Figure 2).

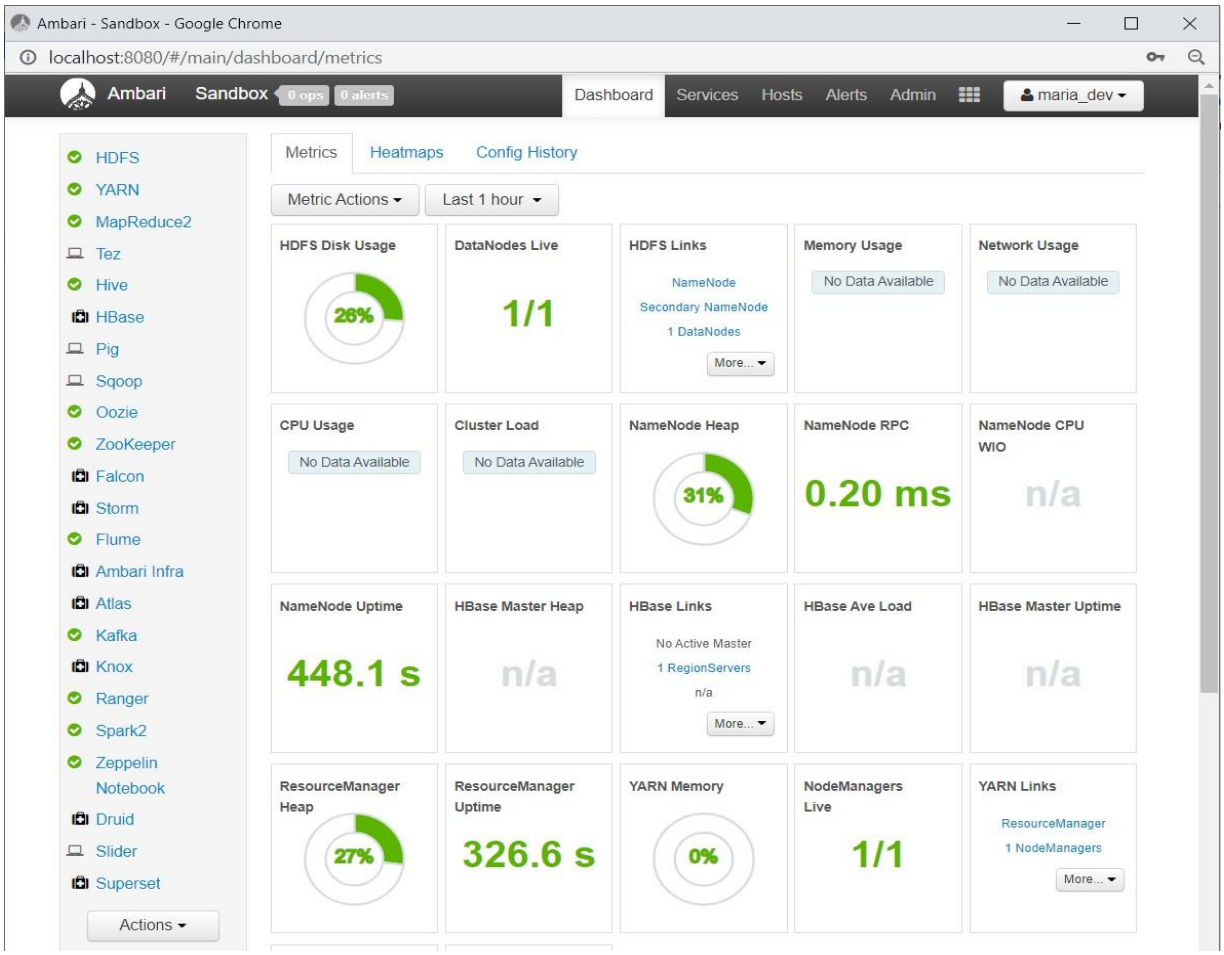

Figure 1. Hortonworks Data Platform

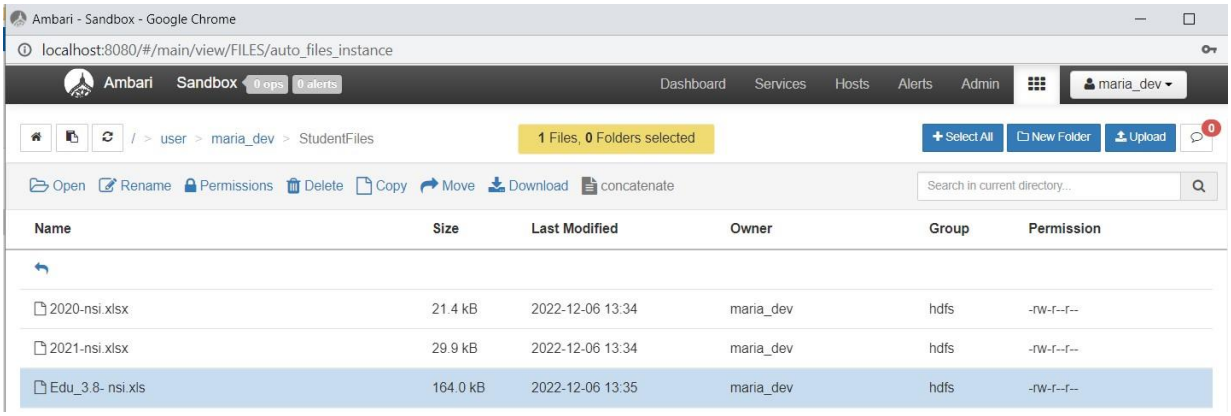

Figure 2. Example of the file storage in Hortonworks Data Platform

The uploaded files can be analyzed using data science methods in two ways: by connection to the storage from data mining software, or by downloading necessary files and analyzing them. ICA is applied using ICrAData software [16] over the dataset for Higher schools, teaching staff, students enrolled and graduates by statistical zones, statistical regions and districts of Bulgaria for 2021, [31]. The input data is presented in the Table 1.

Table 1. The dataset for Higher schools, teaching staff, students enrolled and graduates by statistical zones, statistical regions and districts, 2021, [31]

		Teachers in colleges	Teachers in universities	Enrolled students in colleges	Enrolled students in bachelor	Enrolled students in master	Graduated students in colleges	Graduated students in bachelor	Graduated students in master
		O1	O2	O3	O4	O5	O6	O7	O8
Vidin	C01	0	9	0	453	0	0	69	0
Vratza	C02	0	187	245	978	0	33	151	0
Lovech	C03	31	0	152	0	0	40	0	0
Pleven	C04	60	515	444	845	2 388	106	193	213
V.Tarnovo	C05	0	1 149	0	11 192	4 533	0	2 401	2 203
Gabrovo	C06	0	187	0	3 804	988	0	427	222
Razgrad	C07	0	39	0	208	42	0	40	26
Rousse	C08	0	373	0	4 339	1 393	0	761	529
Silistra	C09	0	78	0	264	0	0	32	0
Varna	C10	25	1 892	1 118	15 140	7 350	381	2 423	1 783
Dobrich	C11	38	0	678	0	0	171	0	0
Shumen	C12	0	538	0	4 392	2 502	0	708	1 303
Bourgas	C13	41	574	601	3 953	2 431	131	612	564
Sliven	C14	35	176	234	693	591	69	100	385
Stara Zagora	C15	93	505	427	2 626	2 509	101	470	481
Yambol	C16	0	64	0	593	78	0	133	56
Blagoevgrad	C17	33	755	524	6 024	2 154	118	1 315	984
Pernik	C18	0	30	0	48	5	0	0	3
Sofia	C19	0	36	115	1 112	306	17	287	147
Sofia City	C20	269	9 949	2 597	57 101	33 892	544	8 922	8 443
Kardzhali	C21	0	154	0	577	0	0	104	0
Plovdiv	C22	47	2 584	664	22 712	12 069	153	3 638	3 469
Smolyan	C23	0	194	0	1 168	270	0	310	139
Haskovo	C24	0	35	0	778	139	0	173	52

The results of the application of the ICA over the dataset for “Education and Lifelong Learning” for 2021 are presented on the Figures 3 and 4, where the degrees of membership of the intercriteria pairs formed within the set of towns are visualized on the Figure 3, and the degrees of non-membership of the intercriteria pairs formed within the set of towns are shown on the next Figure 4.

According to the degrees of membership and degrees of non-membership the intuitionistic fuzzy pairs are classified in the range of strong positive consonance to the strong negative consonance (Table 2).

μ-table	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24
C01	1.00	0.64	0.04	0.54	0.36	0.39	0.68	0.57	0.96	0.11	0.04	0.75	0.36	0.39	0.32	0.71	0.14	0.54	0.25	0.11	0.96	0.04	0.46	0.68
C02	0.64	1.00	0.36	0.79	0.11	0.18	0.46	0.29	0.68	0.14	0.36	0.39	0.54	0.61	0.46	0.50	0.36	0.43	0.32	0.14	0.68	0.36	0.25	0.46
C03	0.04	0.36	1.00	0.29	0.36	0.00	0.00	0.04	0.04	0.75	1.00	0.11	0.43	0.18	0.36	0.00	0.54	0.00	0.21	0.75	0.04	1.00	0.00	0.00
C04	0.54	0.79	0.29	1.00	0.00	0.25	0.54	0.36	0.57	0.11	0.29	0.29	0.61	0.68	0.68	0.57	0.57	0.57	0.32	0.11	0.57	0.29	0.32	0.54
C05	0.36	0.11	0.36	0.00	1.00	0.21	0.11	0.36	0.36	0.54	0.36	0.54	0.04	0.00	0.04	0.11	0.11	0.21	0.04	0.54	0.36	0.36	0.21	0.11
C06	0.39	0.18	0.00	0.25	0.21	1.00	0.71	0.75	0.36	0.04	0.00	0.57	0.43	0.50	0.43	0.68	0.32	0.50	0.79	0.04	0.36	0.00	0.93	0.71
C07	0.68	0.46	0.00	0.54	0.11	0.71	1.00	0.50	0.64	0.00	0.00	0.43	0.36	0.71	0.36	0.96	0.21	0.79	0.57	0.00	0.64	0.00	0.71	0.93
C08	0.57	0.29	0.04	0.36	0.36	0.75	0.50	1.00	0.54	0.11	0.04	0.79	0.57	0.29	0.57	0.54	0.43	0.36	0.54	0.11	0.54	0.04	0.75	0.57
C09	0.96	0.68	0.04	0.57	0.36	0.36	0.64	0.54	1.00	0.11	0.04	0.71	0.32	0.43	0.36	0.68	0.18	0.57	0.21	0.11	1.00	0.04	0.43	0.64
C10	0.11	0.14	0.75	0.11	0.54	0.04	0.00	0.11	0.11	1.00	0.75	0.21	0.21	0.04	0.21	0.00	0.36	0.04	0.07	1.00	0.11	0.75	0.04	0.00
C11	0.04	0.36	1.00	0.29	0.36	0.00	0.00	0.04	0.04	0.75	1.00	0.11	0.43	0.18	0.36	0.00	0.54	0.00	0.21	0.75	0.04	1.00	0.00	0.00
C12	0.75	0.39	0.11	0.29	0.54	0.57	0.43	0.79	0.71	0.21	0.11	1.00	0.46	0.14	0.43	0.46	0.21	0.29	0.36	0.21	0.71	0.11	0.64	0.43
C13	0.36	0.54	0.43	0.61	0.04	0.43	0.36	0.57	0.32	0.21	0.43	0.46	1.00	0.36	0.86	0.39	0.64	0.21	0.64	0.21	0.32	0.43	0.50	0.36
C14	0.39	0.61	0.18	0.68	0.00	0.50	0.71	0.29	0.43	0.04	0.18	0.14	0.36	1.00	0.43	0.68	0.43	0.68	0.64	0.04	0.43	0.18	0.43	0.71
C15	0.32	0.46	0.36	0.68	0.04	0.43	0.36	0.57	0.36	0.21	0.36	0.43	0.86	0.43	1.00	0.39	0.79	0.36	0.57	0.21	0.36	0.36	0.50	0.36
C16	0.71	0.50	0.00	0.57	0.11	0.68	0.96	0.54	0.68	0.00	0.00	0.46	0.39	0.68	0.39	1.00	0.21	0.75	0.54	0.00	0.68	0.00	0.75	0.96
C17	0.14	0.36	0.54	0.57	0.11	0.32	0.21	0.43	0.18	0.36	0.54	0.21	0.64	0.43	0.79	0.21	1.00	0.32	0.46	0.36	0.18	0.54	0.29	0.25
C18	0.54	0.43	0.00	0.57	0.21	0.50	0.79	0.36	0.57	0.04	0.00	0.29	0.21	0.68	0.36	0.75	0.32	1.00	0.36	0.04	0.57	0.00	0.50	0.71
C19	0.25	0.32	0.21	0.32	0.04	0.79	0.57	0.54	0.21	0.07	0.21	0.36	0.64	0.64	0.57	0.54	0.46	0.36	1.00	0.07	0.21	0.21	0.71	0.57
C20	0.11	0.14	0.75	0.11	0.54	0.04	0.00	0.11	0.11	1.00	0.75	0.21	0.21	0.04	0.21	0.00	0.36	0.04	0.07	1.00	0.11	0.75	0.04	0.00
C21	0.96	0.68	0.04	0.57	0.36	0.36	0.64	0.54	1.00	0.11	0.04	0.71	0.32	0.43	0.36	0.68	0.18	0.57	0.21	0.11	1.00	0.04	0.43	0.64
C22	0.04	0.36	1.00	0.29	0.36	0.00	0.00	0.04	0.04	0.75	1.00	0.11	0.43	0.18	0.36	0.00	0.54	0.00	0.21	0.75	0.04	1.00	0.00	0.00
C23	0.46	0.25	0.00	0.32	0.21	0.93	0.71	0.75	0.43	0.04	0.00	0.64	0.50	0.43	0.50	0.75	0.29	0.50	0.71	0.04	0.43	0.00	1.00	0.71
C24	0.68	0.46	0.00	0.54	0.11	0.71	0.93	0.57	0.64	0.00	0.00	0.43	0.36	0.71	0.36	0.96	0.25	0.71	0.57	0.00	0.64	0.00	0.71	1.00

Figure 3. Results of ICA: degrees of membership (μ-table)

ν-table	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24
C01	0.00	0.11	0.32	0.11	0.00	0.25	0.07	0.14	0.04	0.21	0.32	0.07	0.25	0.25	0.29	0.04	0.39	0.11	0.36	0.21	0.04	0.32	0.18	0.07
C02	0.11	0.00	0.25	0.11	0.00	0.50	0.32	0.32	0.07	0.29	0.25	0.18	0.32	0.29	0.39	0.29	0.43	0.25	0.54	0.29	0.07	0.25	0.43	0.32
C03	0.32	0.25	0.00	0.36	0.00	0.43	0.54	0.32	0.32	0.07	0.00	0.21	0.25	0.46	0.32	0.54	0.21	0.43	0.39	0.07	0.32	0.00	0.43	0.54
C04	0.11	0.11	0.36	0.00	0.00	0.54	0.36	0.29	0.07	0.36	0.36	0.18	0.36	0.32	0.29	0.32	0.32	0.21	0.64	0.36	0.07	0.36	0.46	0.36
C05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C06	0.25	0.50	0.43	0.54	0.00	0.00	0.18	0.11	0.29	0.29	0.43	0.11	0.32	0.29	0.32	0.21	0.36	0.29	0.04	0.29	0.29	0.43	0.07	0.18
C07	0.07	0.32	0.54	0.36	0.00	0.18	0.00	0.25	0.11	0.36	0.54	0.14	0.50	0.18	0.50	0.04	0.57	0.11	0.29	0.36	0.11	0.54	0.18	0.07
C08	0.14	0.32	0.32	0.29	0.00	0.11	0.25	0.00	0.18	0.21	0.32	0.04	0.11	0.36	0.11	0.21	0.18	0.29	0.14	0.21	0.18	0.32	0.11	0.18
C09	0.04	0.07	0.32	0.07	0.00	0.29	0.11	0.18	0.00	0.21	0.32	0.11	0.29	0.21	0.25	0.07	0.36	0.07	0.39	0.21	0.00	0.32	0.21	0.11
C10	0.21	0.29	0.07	0.36	0.00	0.29	0.36	0.21	0.21	0.00	0.07	0.14	0.29	0.43	0.29	0.36	0.21	0.29	0.36	0.00	0.21	0.07	0.29	0.36
C11	0.32	0.25	0.00	0.36	0.00	0.43	0.54	0.32	0.32	0.07	0.00	0.21	0.25	0.46	0.32	0.54	0.21	0.43	0.39	0.07	0.32	0.00	0.43	0.54
C12	0.07	0.18	0.21	0.18	0.00	0.11	0.14	0.04	0.11	0.14	0.21	0.00	0.04	0.32	0.07	0.11	0.21	0.18	0.14	0.14	0.11	0.21	0.04	0.14
C13	0.25	0.32	0.25	0.36	0.00	0.32	0.50	0.11	0.29	0.29	0.25	0.04	0.00	0.61	0.14	0.46	0.29	0.54	0.29	0.29	0.29	0.25	0.25	0.50
C14	0.25	0.29	0.46	0.32	0.00	0.29	0.18	0.36	0.21	0.43	0.46	0.32	0.61	0.00	0.54	0.21	0.46	0.11	0.32	0.43	0.21	0.46	0.36	0.18
C15	0.29	0.39	0.32	0.29	0.00	0.32	0.50	0.11	0.25	0.29	0.32	0.07	0.14	0.54	0.00	0.46	0.14	0.39	0.36	0.29	0.25	0.32	0.25	0.50
C16	0.04	0.29	0.54	0.32	0.00	0.21	0.04	0.21	0.07	0.36	0.54	0.11	0.46	0.21	0.46	0.00	0.57	0.14	0.32	0.36	0.07	0.54	0.14	0.04
C17	0.39	0.43	0.21	0.32	0.00	0.36	0.57	0.18	0.36	0.21	0.21	0.21	0.29	0.46	0.14	0.57	0.00	0.36	0.39	0.21	0.36	0.21	0.39	0.54
C18	0.11	0.25	0.43	0.21	0.00	0.29	0.11	0.29	0.07	0.29	0.43	0.18	0.54	0.11	0.39	0.14	0.36	0.00	0.39	0.29	0.07	0.43	0.29	0.18
C19	0.36	0.54	0.39	0.64	0.00	0.04	0.29	0.14	0.39	0.36	0.39	0.14	0.29	0.32	0.36	0.32	0.39	0.39	0.00	0.36	0.39	0.39	0.11	0.29
C20	0.21	0.29	0.07	0.36	0.00	0.29	0.36	0.21	0.21	0.00	0.07	0.14	0.29	0.43	0.29	0.36	0.21	0.29	0.36	0.00	0.21	0.07	0.29	0.36
C21	0.04	0.07	0.32	0.07	0.00	0.29	0.11	0.18	0.00	0.21	0.32	0.11	0.29	0.21	0.25	0.07	0.36	0.07	0.39	0.21	0.00	0.32	0.21	0.11
C22	0.32	0.25	0.00	0.36	0.00	0.43	0.54	0.32	0.32	0.07	0.00	0.21	0.25	0.46	0.32	0.54	0.21	0.43	0.39	0.07	0.32	0.00	0.43	0.54
C23	0.18	0.43	0.43	0.46	0.00	0.07	0.18	0.11	0.21	0.29	0.43	0.04	0.25	0.36	0.25	0.14	0.39	0.29	0.11	0.29	0.21	0.43	0.00	0.18
C24	0.07	0.32	0.54	0.36	0.00	0.18	0.07	0.18	0.11	0.36	0.54	0.14	0.50	0.18	0.50	0.04	0.54	0.18	0.29	0.36	0.11	0.54	0.18	0.00

Figure 4. Results of ICA: degrees of non-membership (ν-table)

The pairs of towns in strong positive consonance and positive consonance in the field of "Education and Lifelong Learning" are shown below. Pairs of criteria in strong positive consonance [0,95; 1,00] are: Vidin – Silistra, Lovech – Dobrich, Razgrad – Yambol, Varna – Sofia City, Silistra – Kurdzhali, Dobrich – Plovdiv, Lovech – Plovdiv, Vidin – Kurdzhali.

Pairs of criteria in positive consonance [0,85; 0,95) have the following form: Bourgas – Stara Zagora, Gabrovo – Smolyan, Razgrad – Haskovo.

Table 2. Number of ICA pairs of criteria per type of correlation

Type of correlation	Number of pairs of criteria for a year 2021
strong positive consonance [0,95; 1,00]	8
positive consonance [0,85; 0,95)	3
weak positive consonance [0,75; 0,85)	16
weak dissonance [0,67; 0,75)	22
dissonance [0,57; 0,67)	24
strong dissonance [0,43; 0,57)	52
dissonance [0,33; 0,43)	36
weak dissonance [0,25; 0,33)	20
weak negative consonance [0,15;0,25)	24
negative consonance [0,05;0,15)	24
strong negative consonance [0,00;0,05]	49

In order to determine the tendency between enrolled and graduated students and teachers in the towns of Bulgaria we observe the results of the application of InterCriteria Analysis. If two towns are in strong correlation, we can claim the dependencies between them. If two towns are in dissonance, we can claim that they are independent according to the students behavior. If two towns are in negative consonance, we can claim that they have opposite behavior according to the students education. The results in strong positive consonance and positive consonance present Bulgarian towns with similar student migration. The obtained results are interpreted in the intuitionistic fuzzy triangle (Figure 5).

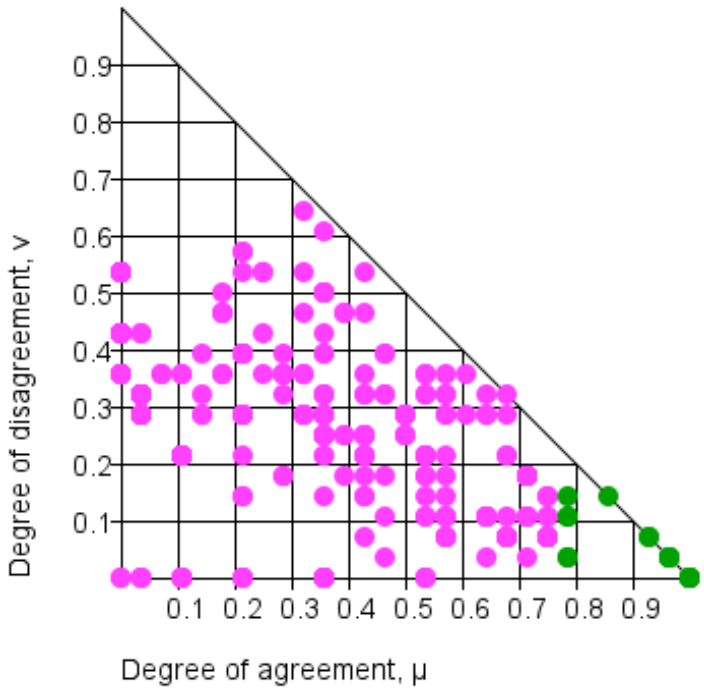

Figure 5. ICA results plotted onto the intuitionistic fuzzy interpretational triangle

The ICA is applied to a transposed version of the dataset, to determine the relations between enrolled and graduated students and teachers. The degrees of membership and degrees of non-membership are presented in the Figure 6.

μ-table	O1	O2	O3	O4	O5	O6	O7	O8
O1	1.00	0.30	0.72	0.18	0.13	0.84	0.22	0.36
O2	0.30	1.00	0.39	0.43	0.36	0.33	0.78	0.71
O3	0.72	0.39	1.00	0.26	0.20	0.83	0.30	0.43
O4	0.18	0.43	0.26	1.00	0.47	0.19	0.45	0.33
O5	0.13	0.36	0.20	0.47	1.00	0.14	0.43	0.62
O6	0.84	0.33	0.83	0.19	0.14	1.00	0.23	0.36
O7	0.22	0.78	0.30	0.45	0.43	0.23	1.00	0.63
O8	0.36	0.71	0.43	0.33	0.62	0.36	0.63	1.00

ν-table	O1	O2	O3	O4	O5	O6	O7	O8
O1	0.00	0.32	0.09	0.27	0.25	0.07	0.36	0.22
O2	0.32	0.00	0.24	0.29	0.26	0.38	0.17	0.14
O3	0.09	0.24	0.00	0.24	0.24	0.08	0.31	0.19
O4	0.27	0.29	0.24	0.00	0.16	0.32	0.27	0.29
O5	0.25	0.26	0.24	0.16	0.00	0.29	0.19	0.14
O6	0.07	0.38	0.08	0.32	0.29	0.00	0.43	0.28
O7	0.36	0.17	0.31	0.27	0.19	0.43	0.00	0.17
O8	0.22	0.14	0.19	0.29	0.14	0.28	0.17	0.00

Figure 6. Results of ICA on the transposed dataset:
degrees of membership (μ-table, left) and non-membership(ν-table, right)

According to the degrees of membership and degrees of non-membership the intuitionistic fuzzy pairs are classified in the range of strong positive consonance to the strong negative consonance (Table 3).

Table 3. Number of intuitionistic fuzzy pairs – ICA on the transposed dataset

Type of correlation	Number of pairs of criteria
strong positive consonance [0,95; 1,00]	
positive consonance [0,85; 0,95)	
weak positive consonance [0,75; 0,85)	3
weak dissonance [0,67; 0,75)	4
dissonance [0,57; 0,67)	
strong dissonance [0,43; 0,57)	5
dissonance [0,33; 0,43)	6
weak dissonance [0,25; 0,33)	3
weak negative consonance [0,15;0,25)	5
negative consonance [0,05;0,15)	2
strong negative consonance [0,00;0,05]	

The pairs of educated people in weak positive consonance in 2021 in the field of higher education in Buglaria are shown below. Pairs of criteria in weak positive consonance [0,95; 1,00] are: Teachers in colleges – Graduated students in colleges; Teachers in universities – Graduated students in bachelor; Enrolled students in colleges – Graduated students in colleges.

In order to determine the tendency between enrolled and graduated students and teachers in the towns of Bulgaria we should observe the results of the application of InterCriteria Analysis. If the two groups of educated people are in strong correlation, we can claim dependencies between them. If the two groups of educated are in dissonance, we can claim that they are independent according to the students behavior. If the two groups of educated are in negative consonance, we can claim that they have opposite behavior according to the students education.

Thereafter the results present weak relationships between the pairs “Teachers in colleges – Graduated students in colleges”, “Teachers in universities – Graduated students in bachelor”, “Enrolled students in colleges – Graduated students in colleges”. The obtained results are interpreted in the intuitionistic fuzzy triangle (Figure 7).

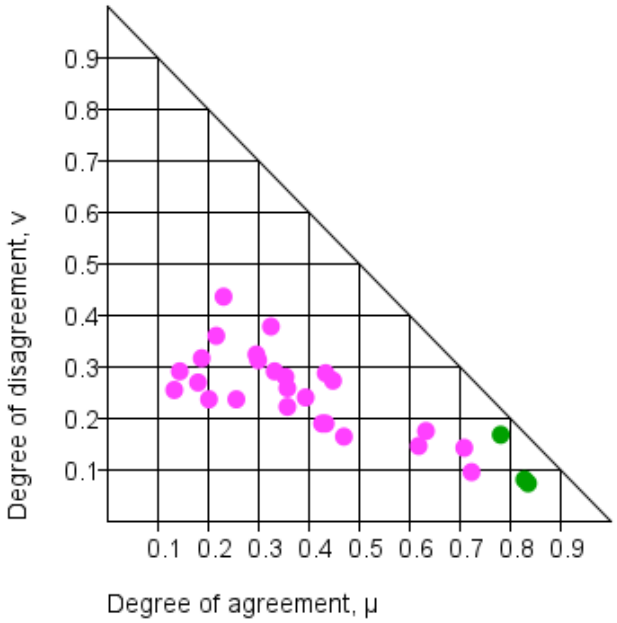

Figure 7. Results of ICA on the transposed dataset,
plotted onto the intuitionistic fuzzy interpretational triangle

The application of the ICA method over datasets for Higher schools, teaching staff, students enrolled and graduates by statistical zones, statistical regions and districts in the Republic of Bulgaria can help to identify the most correlated Bulgarian towns according to enrolled and graduated students and teachers, determine the tendencies between enrolled and graduated students and help to monitor the behavior of the objects or criteria.

3 Conclusion

In the recent years, Big Data systems have become very popular. The traditional data warehouses are transformed into data lakes which allows data storage in the form of different types of files. Thereafter data lakehouses are introduced to combine the relational databases with NoSQL databases (files-based databases). Big Data storages are typically used for applying procedures for decision making, data mining, reporting and OLAP analyses. In the current investigation InterCriteria Analysis (ICA) as a method allowing decision making on the base of the discovered correlations between the data is successfully applied in Big Data field to a dataset regarding Bulgarian national system of higher education in 2021, and outlining dependencies and recent trends.

Acknowledgements

The present research has been supported by the Bulgarian National Science Fund under Grant Ref. No. KP-06-N-22/1 “Theoretical Research and Applications of InterCriteria Analysis”.

References

[1] Atanassov, K., Atanassova, V., & Gluhchev, G. (2015). InterCriteria Analysis: Ideas and Problems. *Notes on Intuitionistic Fuzzy Sets*, 21(1), 81–88.

[2] Atanassov, K., & Bureva, V. (2022). Two new operations over extended index matrices and their applications in Big Data. *Proceedings of 17th Conference on Computer Science and Intelligence Systems (FedCSIS)*, 2022, 1–6.

[3] Atanassov, K., Mavrov, D., & Atanassova, V. (2014). Intercriteria Decision Making: A New Approach for Multicriteria Decision Making, Based on Index Matrices and Intuitionistic Fuzzy Sets. *Issues in Intuitionistic Fuzzy Sets and Generalized Nets*, 11, 1–8.

[4] Atanassova, V., Doukovska, L., Michalikova, A., & Radeva, I. (2016). Intercriteria analysis: From pairs to triples. *Notes on Intuitionistic Fuzzy Sets*, 22(5), 98–110.

[5] Atanassova, V., Mavrov, D., Doukovska, L., & Atanassov, K. (2014). Discussion on the threshold values in the InterCriteria Decision Making approach. *Notes on Intuitionistic Fuzzy Sets*, 20(2), 94–99.

[6] Atanassova, V., & Vardeva, I. (2014). Sum- and average-based approach to criteria shortlisting in the InterCriteria Analysis. *Notes on Intuitionistic Fuzzy Sets*, 20(4), 41–46.

[7] Atanassova, V., Vardeva, I., Sotirova, E., & Doukovska, L. (2016). Traversing and ranking of elements of an intuitionistic fuzzy set in the intuitionistic fuzzy interpretation triangle, Chapter, Novel Developments in Uncertainty Representation and Processing, Vol. 401, Advances in Intelligent Systems and Computing, 161–174.

[8] Bureva, V. (2019/2020). Index Matrices as a Tool for Data Lakehouse Modelling. *Annual of “Informatics” Section of the Union of Scientists in Bulgaria*, 10, 81–105.

[9] Bureva, V. (2021). Overview of Big Data Systems. *Annual of “Informatics” Section of the Union of Scientists in Bulgaria*, 11, 1–14.

[10] Bureva, V., Michalíková, A., Sotirova, E., Popov, S., Riečan, B., & Roeva, O. (2017). Application of the InterCriteria Analysis to the universities rankings system in the Slovak Republic. *Notes on Intuitionistic Fuzzy Sets*, 23(2), 128–140.

[11] Bureva, V., Sotirova, E., & Panayotov, H. (2015–2016). The intercriteria decision making method to Bulgarian university ranking system. *Annual of “Informatics” Section, Union of Scientists in Bulgaria*, 8, 54–70.

[12] Bureva, V., Sotirova, E., Sotirov, S., & Mavrov, D. (2015). Application of the InterCriteria decision making method to Bulgarian universities ranking. *Notes on Intuitionistic Fuzzy Sets*, 21(2), 111–117.

[13] Chorukova, E., Marinov, P., & Umlenski, I. (2021). Survey on Theory and Applications of InterCriteria Analysis Approach. In: Atanassov K.T. (eds) *Research in Computer Science in the Bulgarian Academy of Sciences*. Studies in Computational Intelligence, Vol. 934, pp. 453–469. Springer, Cham.

[14] Çuvalcıoğlu, G., Bureva, V., & Michalikova, A. (2019). Intercriteria analysis applied to university ranking system of Turkey. *Notes on Intuitionistic Fuzzy Sets*, 25(4), 90–97.

[15] Doukovska, L., Atanassova, V., Sotirova, E., Vardeva, I., & Radeva, I. (2019). Defining consonance thresholds in InterCriteria analysis: An overview. In: Hadjiski, M., & Atanassov, K.T. (eds.). *Intuitionistic Fuzziness and Other Intelligent Theories and Their Applications*. Studies in Computational Intelligence, 757, 161–179. Springer, Cham.

[16] Ikonomov, N., Vassilev, P., & Roeva, O. (2018). ICrAData – Software for InterCriteria Analysis. *International Journal Bioautomation*, 22(1), 1–10.

[17] InterCriteria Analysis Portal (2022). Publications. Available online at: `https://intercriteria.net/publications/`

[18] InterCriteria Analysis Portal (2022). Software. Available online at: `https://intercriteria.net/software/`

[19] Krawczak, M., Bureva, V., Sotirova, E., & Szmidt, E. (2016). Application of the InterCriteria Decision Making Method to Universities Ranking. *Novel Developments in Uncertainty Representation and Processing*. Advances in Intelligent Systems and Computing Series, 401, 365–372, Springer, Cham.

[20] Mavrov, D. (2015). Software for InterCriteria Analysis: Implementation of the main algorithm. *Notes on Intuitionistic Fuzzy Sets*, 21(2), 77–86.

[21] Mavrov, D. (2015-2016). Software for intercriteria analysis: Working with the results. *Annual of "Informatics" Section, Union of Scientists in Bulgaria*, 8, 37–44.

[22] Orozova, D., & Popchev, I. (2020). Cyber-Physical-Social Systems for Big Data. *Proceedings of the 21st Int. Symposium on Electrical Apparatus and Technologies (SiELA)*, 3-6 June 2020, Burgas, 334-337, DOI: 10.1109/SIELA49118.2020.9167161

[23] Parvathi, R., Atanassova, V., Doukovska, L., Yuvapriya, C., & Indhurekha, K. (2018). InterCriteria Analysis of rankings of Indian universities. *Notes on Intuitionistic Fuzzy Sets*, 24 (1), 99–109.

[24] Popchev, I. & Georgieva, M. (2015). Sciencemetric Analysis in the "Big Data" Area. *Scientific conference with International Participation "Horizons in the Development of Human Resources and Knowledge"*, 12-14 June 2015, Burgas, Vol. 2, 469-476.

[25] Popchev, I., & Orozova, D. (2019). Towards Big Data Analytics in the e-Learning Space. *Cybernetics and Information Technologies*, 19(3), 16-24.

[26] Popchev, I., Orozova, D., & Stoyanov, S. (2019). IoT and Big Data Analytics in E-Learning. *Proceedings of the 2019 IEEE Conference on Big Data, Knowledge and Control Systems Engineering (BdKCSE)*, 21-22 November 2019, Sofia, pp. 1-5, DOI: 10.1109/BdKCSE48644.2019.9010666.

[27] Roeva, O., Vassilev, P., Ikonomov, N., Angelova, M., Su, J., & Pencheva, T. (2019). On Different Algorithms for InterCriteria Relations Calculation. Hadjiski, M., Atanassov, K. (eds). *Intuitionistic Fuzziness and Other Intelligent Theories and Their Applications.* Studies in Computational Intelligence, 757, 143–160. Springer, Cham.

[28] Sotirova, E., & Shannon, A. (2015). Application of InterCriteria Decision Making to the rankings of Australian universities. *Notes on Intuitionistic Fuzzy Sets*, 21(4), 136–142.

[29] Sotirova, E., Bureva, V., Chountas, P., & Krawczak, M. (2016). An application of intercriteria decision making method to the rankings of universities in the United Kingdom. *Notes on Intuitionistic Fuzzy Sets*, 22(3), 112–119.

[30] Traneva, V., Tranev S., Szmidt E., Atanassov K. (2018) Three Dimensional Intercriteria Analysis over Intuitionistic Fuzzy Data. In: Kacprzyk J., Szmidt E., Zadrożny S., Atanassov K., Krawczak M. (eds) Advances in Fuzzy Logic and Technology 2017. EUSFLAT 2017, IWIFSGN 2017. Advances in Intelligent Systems and Computing, vol 643, pp. 442-449. Springer, Cham.

[31] Tsankova, E. (2022). Higher schools, teaching staff, students enrolled and graduates by statistical zones, statistical regions and districts. *National Statistics Institute of the Republic of Bulgaria*. Available online at: `https://nsi.bg/sites/default/files/files/data/timeseries/Edu_3.8_en.xls`