



Министерство науки и высшего образования Российской Федерации
Калужский филиал
федерального государственного бюджетного
образовательного учреждения высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(КФ МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИУК «Информатика и управление»

КАФЕДРА ИУК4 «Программная инженерия»

Лабораторная работа №3

«Технологии обработки данных в системе KNIME Analytics Platform»

**ДИСЦИПЛИНА: «Интеллектуальные информационные системы
анализа данных»**

Выполнил: студент гр. ИУК4-21М _____ (Сафронов Н.С.)
(подпись) (Ф.И.О.)

Проверил: _____ (Белов Ю.С.)
(подпись) (Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга, 2025

Цель работы: формирование практических навыков работы в программе KNIME.

Задание:

1. Используя набор данных согласно варианту, реализовать модель машинного обучения в соответствии с примером 2, визуализировать полученные результаты.

2. Подготовить отчёт.

Вариант 7

Water Quality

Результат выполнения работы

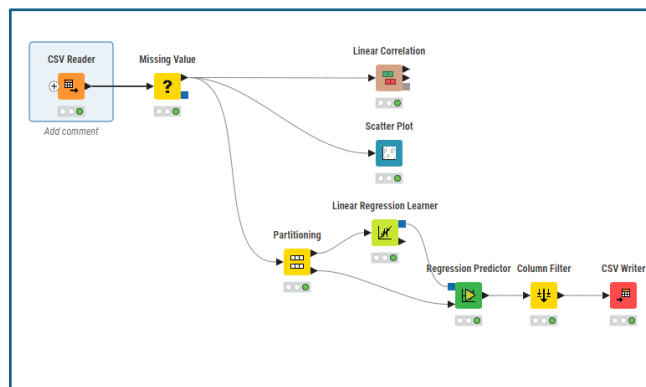


Рисунок 1 – Архитектура решения

The screenshot shows the 'CSV Reader' dialog box in KNIME. The 'File' field contains the path 'E:\Dev\bmcs-magistracy\2nd-term\intelligent-data-analysis\lab5\data\water_potability.csv'. The 'Reader options' section shows 'Autodetect format' selected. The 'Preview' section displays a table with 22 rows and 12 columns.

Row ID	pH	D Hardness	D Solids	D Chloro...	D Sulfate	D Conduct...	D Organic...	D Trihalo...	D Turbidity	T Potability
Row0	7	204.89	25.761	3.29	7.3	368.538	964.309	10.38	86.991	2.963
Row1	7	129.423	18.432	0.88	6.433	362.883	15.18	96.329	4.931	0
Row2	8.099	224.238	19.909	5.42	9.278	7	418.606	16.869	66.42	3.056
Row3	8.317	214.373	22.238	4.07	9.039	166.886	263.267	16.437	109.242	4.629
Row4	9.492	181.102	17.978	9.86	6.547	310.136	398.411	11.058	31.998	4.079
Row5	5.304	188.313	28.748	6.88	7.545	325.678	280.468	8.4	24.528	2.56
Row6	10.324	248.072	28.769	7.27	7.513	293.663	383.652	13.79	84.604	2.973
Row7	8.436	203.362	13.672	0.82	4.363	303.31	434.608	13.364	62.768	4.401
Row8	7	138.869	14.265	5.84	7.804	268.647	389.376	12.706	53.629	3.995
Row9	11.18	227.231	25.484	5.08	9.077	404.042	563.885	17.928	71.937	4.371
Row10	7.361	185.521	32.452	6.14	7.551	326.624	425.383	15.387	76.74	3.662
Row11	7.975	238.693	18.767	6.67	8.11	364.098	14.526	76.486	4.912	0
Row12	7.12	156.705	18.730	8.14	3.408	282.344	347.713	15.83	79.501	3.446
Row13	7	150.175	27.231	3.62	6.838	295.436	379.762	19.371	76.51	4.414
Row14	7.496	205.345	28.388	0.05	5.073	7	444.645	13.228	76.3	4.777
Row15	6.347	186.723	41.083	2.39	9.63	364.468	336.743	11.04	79.072	4.376
Row16	7.652	211.049	26.988	6.61	10.895	7	315.145	26.797	86.652	4.368
Row17	9.182	273.814	24.041	3.28	6.903	398.351	477.979	13.387	71.457	4.504
Row18	8.915	276.357	19.465	3.88	5.204	7	451.444	12.889	63.821	2.436
Row19	7.371	214.467	25.630	32	4.433	335.754	469.913	12.909	62.767	2.56
Row20	7	227.435	22.305	5.67	10.334	7	194.82	16.332	45.383	4.133
Row21	6.46	188.384	30.546	3.64	6.89	310.931	515.873	17.884	77.042	5.75
Row22	7	215.978	17.107	2.24	5.607	326.944	436.256	14.189	59.855	5.459

Рисунок 2 – Загруженный файл

#	RowID	ph Number (double)	Hardness Number (double)	Solids Number (double)	Chloramines Number (double)	Sulfate Number (double)	Conductivity Number (double)	Organic_carbon Number (double)	Trihalomethanes Number (double)	Turbidity Number (double)	Potability Number (double)
1	ph	1	0.076	-0.082	-0.032	0.014	0.017	0.04	0.003	-0.036	-0.003
2	Hardne	0.076	1	-0.047	-0.03	-0.093	-0.024	0.004	-0.013	-0.014	-0.014
3	Solids	-0.082	-0.047	1	-0.07	-0.15	0.014	0.01	-0.009	0.02	0.034
4	Chloram	-0.032	-0.03	-0.07	1	0.024	-0.02	-0.013	0.017	0.002	0.024
5	Sulfate	0.014	-0.093	-0.15	0.024	1	-0.014	0.027	-0.026	-0.01	-0.02
6	Condu	0.017	-0.024	0.014	-0.02	-0.014	1	0.021	0.001	0.006	-0.008
7	Organic	0.04	0.004	0.01	-0.013	0.027	0.021	1	-0.013	-0.027	-0.03
8	Trihalo	0.003	-0.013	-0.009	0.017	-0.026	0.001	-0.013	1	-0.021	0.007
9	Turbidi	-0.036	-0.014	0.02	0.002	-0.01	0.006	-0.027	-0.021	1	0.002
10	Potabil	-0.003	-0.014	0.034	0.024	-0.02	-0.008	-0.03	0.007	0.002	1

Рисунок 3 – Матрица корреляции



Рисунок 4 – Точечный график зависимости жёсткости воды от её рН

#	RowID	ph Number (double)	Hardness Number (double)	Solids Number (double)	Chloramines Number (double)	Sulfate Number (double)	Conductivity Number (double)	Organic_car... Number (double)	Trihalometh... Number (double)	Turbidity Number (double)	Potability Number (double)	Prediction (P... Number (double)
1	Row2	8.099	224.236	19,909.542	9.276	333.074	418.606	16.869	66.42	3.056	0	0.4
2	Row6	10.224	248.072	28,749.717	7.513	393.663	283.652	13.79	84.604	2.673	0	0.442
3	Row12	7.12	156.705	18,730.814	3.606	282.344	347.715	15.93	79.501	3.446	0	0.391
4	Row15	6.347	186.733	41,065.235	9.63	364.488	516.743	11.54	75.072	4.376	0	0.451
5	Row19	7.371	214.497	25,630.32	4.433	335.754	469.915	12.509	62.797	2.56	0	0.374
6	Row22	7.037	215.978	17,107.224	5.607	326.944	436.256	14.189	59.855	5.459	0	0.384
7	Row24	5.4	140.739	17,266.593	10.057	328.358	472.874	11.256	56.932	4.825	0	0.421
8	Row27	7.037	145.768	13,224.936	7.906	304.002	298.991	12.73	49.537	4.005	0	0.432
9	Row34	7.037	232.28	14,787.206	5.475	333.074	383.982	12.167	86.081	5.029	0	0.391
10	Row35	5.116	191.953	19,620.545	6.061	323.836	441.748	10.966	49.238	3.902	0	0.386
11	Row38	7.037	143.454	19,942.273	5.891	333.074	427.131	22.47	53.124	2.908	0	0.362
12	Row40	7.037	233.859	11,703.924	4.599	309.039	349.4	18.339	42.677	3.51	0	0.353
13	Row42	7.146	238.69	28,780.34	6.814	385.976	332.033	11.093	66.138	5.183	0	0.438
14	Row43	9.921	202.817	9,973.934	6.882	337.351	333.192	23.918	71.834	4.691	0	0.379
15	Row44	4.758	183.349	21,568.429	4.731	333.074	403.944	18.668	66.912	4.543	0	0.366
16	Row49	8.757	200.191	21,536.225	4.915	317.883	404.718	13.768	47.931	3.626	0	0.397
17	Row51	7.81	100.458	12,013.551	5.212	247.201	605.22	9.611	66.084	2.447	0	0.368

Рисунок 5 – Результат работы предиктора

Выводы: в ходе выполнения работы были сформированы практические навыки работы в программе KNIME.