Міністерство освіти і науки України

Черкаський державний технологічний університет

Кафедра програмного забезпечення автоматизованих систем

Звіт

З лабораторної роботи №4

З предмету «Аналіз великих даних»

Перевірив

доцент к. ПЗАС

Рідкокаша А. А.

Виконав

Студент IV курсу

Група ПЗ-154

Кравченко А. О.

Черкаси, 2019

**Тема:** Знайомство з Weka API для регресійного аналізу.

**Мета:** Ознайомитися та отримати навички практичного застосування методів бібліотеки Weka для регресійної класифікації даних.

**Постановка завдання:** Розробити програму для побудови регресійної моделі даних.

|  |  |
| --- | --- |
| **Варіант** | **Завдання** |
| 13 | LinearRegression |

**Виконання лабораторної роботи:**

Для даної лабораторної роботи було створено файл ICU.arff із наступними атрибутами:

% ID: ID number of the patient

% STA: Vital status (0 = Lived, 1 = Died)

% AGE: Patient's age in years

% SEX: Patient's sex (0 = Male, 1 = Female)

% RACE: Patient's race (1 = White, 2 = Black, 3 = Other)

% SER: Service at ICU admission (0 = Medical, 1 = Surgical)

% CAN: Is cancer part of the present problem? (0 = No, 1 = Yes)

% CRN: History of chronic renal failure (0 = No, 1 = Yes)

% INF: Infection probable at ICU admission (0 = No, 1 = Yes)

% CPR: CPR prior to ICU admission (0 = No, 1 = Yes)

% SYS: Systolic blood pressure at ICU admission (in mm Hg)

% HRA: Heart rate at ICU admission (beats/min)

% PRE: Previous admission to an ICU within 6 months (0 = No, 1 = Yes)

% TYP: Type of admission (0 = Elective, 1 = Emergency)

% FRA: Long bone, multiple, neck, single area, or hip fracture

% PO2: PO2 from initial blood gases (0 = >60, 1 = <60)

% PH: PH from initial blood gases

% BIC: Bicarbonate from initial blood gases

% LOC: Level of consciousness at admission

@RELATION relation

@ATTRIBUTE 'AGE' numeric

@ATTRIBUTE 'SYS' numeric

@ATTRIBUTE 'HRA' numeric

@ATTRIBUTE 'PRE' numeric

@ATTRIBUTE 'TYP' numeric

@ATTRIBUTE 'FRA' numeric

@ATTRIBUTE 'PO2' numeric

@ATTRIBUTE 'PH' numeric

@ATTRIBUTE 'PCO' numeric

@ATTRIBUTE 'BIC' numeric

@ATTRIBUTE 'CRE' numeric

@ATTRIBUTE 'LOC' numeric

На рисунку 1 зображено короткі відомості про завантажені дані та атрибути моделі. Зокрема на рисунку наведено інформацію по атрибуту HRA, що відповідає за показник пульсу в пацієнта.

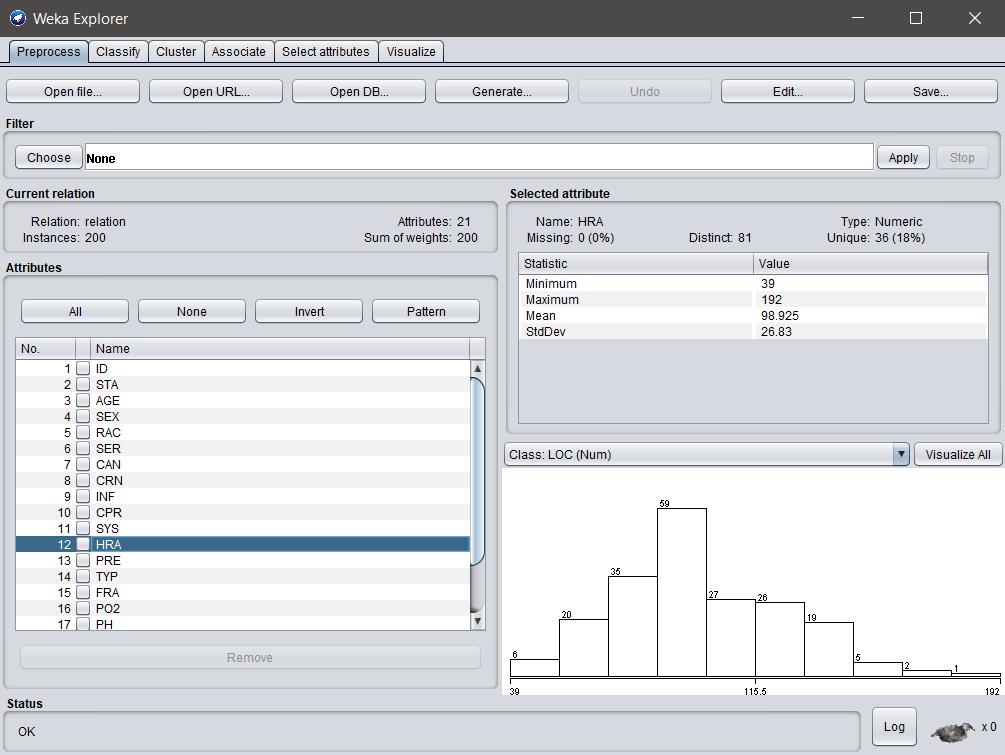


Рис.1 – інформація про ICU.arff

Для прогнозування стану пацієнта за результатами значень його медичного обстеження було використано LinearRegression. На рисунку 2 зображено результат обрахунків та налаштування конфігурації для правильних розрахунків.

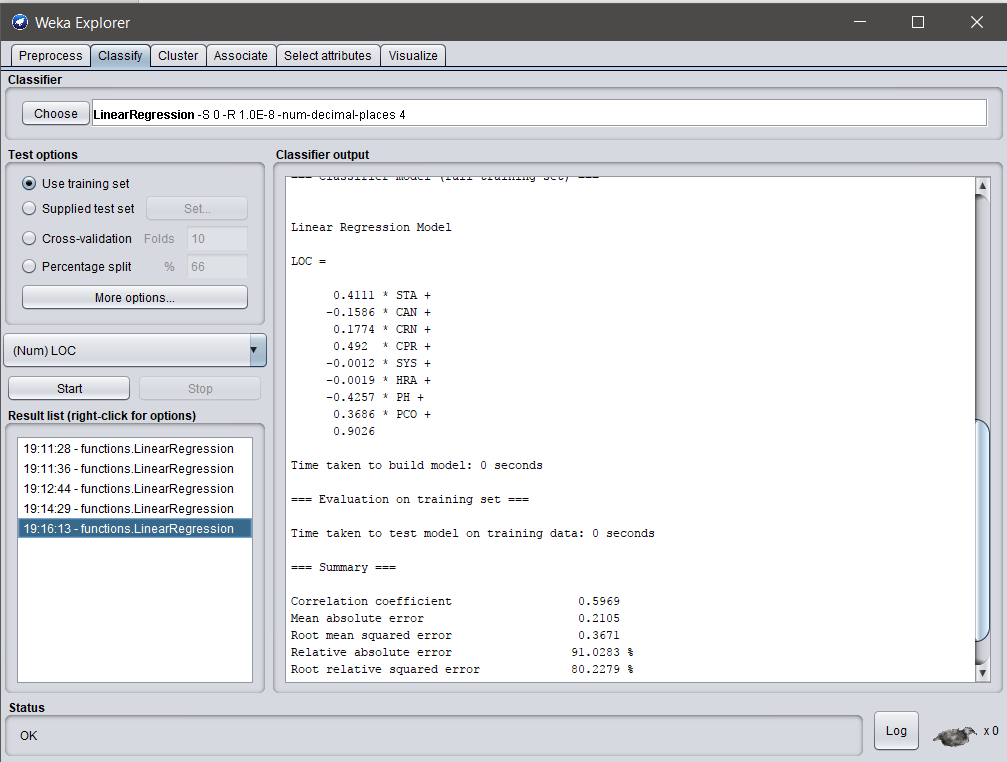


Рис. 2 – результат побудови LinearRegression.

**Висновок:** під час виконання лабораторної роботи, я навчився будувати LinearRegression правила із допомогою середовища Weka. Ознайомився із параметрами побудови та залежним від цього отриманим результатом.