



Patronage 2019

Machine Learning & Big Data

Zadanie 1

Cel podstawowy zadania:

Za pomocą języka Python napisać program który na podstawie danych wejściowych wykona przewidywanie (ang. estimating) zarobków dla danych które nie są kompletne w zbiorze danych wejściowych.

Program powinien wykonać przewidywanie na podstawie danych które są kompletne.

Po obliczeniu prognozy prezentacja wyników za pomocą funkcji print w konsoli.

Cel fakultatywny zadania:

prezentacja graficzna wyniku prognozy za pomocą wykresu przebiegu funkcji y=f(x) gdzie y=predicted_salary_brutto, x=workedYears

Sugerowana metoda prognozy:

Sugeruję wykonanie zadania wykorzystując metodę regresji liniowej (ang. linear regression), niemniej minimalnie niniejsze zadanie można wykonać także wykorzystując prostsze statystyczne metody niż regresja.

(Np. wyliczania median (ang. medians) czy też średnich ucinanych (ang. trimmed mean) dla każdej z brakujących wartości z określonego roku stażu pracy itp.)

Sugerowane biblioteki python:

, pandas, sklearn, matplotlib, numpy,

Sposób przesłania rozwiązania:

Proszę o przesyłanie odpowiedzi (kodu programu) na adres e-mail:

patronage-cognitive-leaders@intive.com jako archiwum zip

Jeżeli ktoś prześle email ze skrótem do github lub Bitbucket będzie to również jak najbardziej zaakceptowane.





Informacja uzupełniająca:

Celem testu jest przede wszystkim zaprezentowanie sposobu podejścia do rozwiązywania przedstawionego zadania za pomocą języka Python przy wykorzystaniu podstawowych metod statystycznych

Krzysztof Kawczyński, Al Software Developer – Cognitive Services krzysztof.kawczynski@intive.com

BLStream, SMT Software Services and Kupferwerk are now INTIVE – learn more on intive.com.

BLStream S.A.. | plac Hołdu Pruskiego 9 | 70-550 Szczecin

District Court for Szczecin-Centrum in Szczecin

XIII Commercial Division of the National Court Register

KRS: 0000652746 | NIP: 8513202115 | Share Capital: PLN 300,000