Anfertigen eines CAT bzw. PIKE

Kann … // Algorithmus x // exact … // (Kunden)-Problem … berechnen / lösen?
Yes, a Neural Network can performe the regression

Big Data / Data Mining / Data Science Kernaussage:
(P) roblem):
[Welcher Frage hat für die Lösung des Kunden / Auftraggeber die größte Bedeutung?]
Medical expense is one of the major recurring expenses in human life. It is aslo known that the life style and variour parameter dominates the health.And subsequently impacts the medical expenses.According to the studies the parameter which contributes the expenses are smoking, age and BMI.In this analysis we study hoe these parameters are correlated between them and predict the expenses by creating a regression model

(I) ntervention:
(Bibliotheken- und Algorithmen-Auswahl, … z.B. pandas für Finanzdaten …)
[Welche Berechnung erwäge ich vornehmlich?]:
For data analysis and preprocessing pandas, numpy, matplotlib,Plotly and seaborn
were used.For building and training the neural network, the Tensorflow library was used
with the Keras high-level interface.Also somefunction from Scikit learn.

(K) ontrollintervention
(falls erforderlich: Bliotheken- und Algorithmen-Auswahl … z.B. scikit-learn für Finanzdaten …)
[Was ist die andere Möglichkeit?]:
Other Neural Network libaries are also availabe like pytorch but Keras is the best when working with small datasets, rapid prototyping, and multiple back-end support.

(E) rgebnismaß (Zielgröße(n)) – Die Evidence
[Was möchte ich / der Kunde erreichen? Z.B. Prädiktor oder Klassifikator erstellen …]:
To do regression analysis and build a model to predict the medical expenses depending on the factors which affect the life expectancy of a customer.

Anmerkungen:
Literaturhinweise: https://www.moneycrashers.com/factors-health-insurance-premium-costs/

Die Suche nach der besten Evidenz

1. **Problem**
    creating an automated system to estimate the annual medical expenditure for new customers, using information such as their age, gender, BMI, children, smoking habits and region of residence.

2. **Definition einer wichtigen suchbaren Frage**
    Does our model will predict the cost accurely with different instances?

3. **Auswahl der wahrscheinlichsten Quelle für diese Evidenz**
    After training, the performance of the model is tested on a selection of the existing test data and evaluated.

4. **Erstellung einer Suchstrategie**
    Different neural network architectures (number of layers, number of units, activation, etc.) are studied to achieve the highest possible accuracy on the test set. In order to avoid overfitting of the training data, early stopping stratergy is used.

5.   **Zusammenstellung der Evidenzausbeute**
A model with an input attributes of 6 and 16 units, with two hidden layers of 16 and 32 units each, and an output layer of 1 unit predicts the cost of medical expenses . And using early stopping to avoid overfitting

6.   **Anwendung der Evidenz**
I have built a model which predicts the cost of medical expenses on the factors like age, gender, region and BMI