

# **Stromerzeugung bei unterschiedlichen Wetterlagen – Datenaufbereitung und Analyse zur Korrelation Mithilfe von Regressionsalgorithmen in Weka**

**Korbinian Eller, Kay Gietenbruch,**

## **Abstract**

In dieser Ausarbeitung wird das Vorgehen beim Sammeln von Daten, deren Aufbereitung und die Analyse mithilfe der Regressionsalgorithmen in Weka erläutert und die Ergebnisse dokumentiert.

Ziel ist es durch Wetterdaten des Deutschen Wetter Dienstes (DWD) auf Stromerzeugungszahlen der erneuerbaren Energien Solar und Wind (größte Beeinflussung durch Wetter) zu schließen

geeignetste Modell für die Analyse der Korrelation ist. So wurde der Plan neu geschrieben und eine Analyse mithilfe von Regressionsalgorithmen in dem Machine-Learning Programm Weka stand ab dem Zeitpunkt im Vordergrund. Dafür sollten die Daten noch angepasst und dann mit Weka und den Regressionsalgorithmen experimentiert werden. In dem Sinne, dass das am best geeignete Modell zu Regression gefunden wird!

## **Daten**

## **Idee**

Zur gegebenen Themenstellung "Einen Themenbereich der KI vertiefen" war eine erste Idee die Klassifikation von Datensätzen. Dies war sehr ähnlich einer der letzten Aufgabenstellungen der Übungsstunden des Fachs. Es war vor allem aufgrund der Implementierung in der Programmiersprache C und damit dem greifbar machen des Algorithmus interessant.

Der Gedanke der Klassifikation war schnell gefestigt nun fehlten noch die Daten mit denen gearbeitet werden soll. Über Daten wie Covid-19 Erkrankungs Zahlen, Bußgeldbescheide, Denkmalstandorte oder Amazon Personen Daten ist eine Idee herausgestochen.

"Es wäre doch interessant, Wetterdaten und Stromdaten in korrelation miteinander zu bringen und so die Stromerzeugung anhand des vorherrschenden Wetters deuten zu können". Und das war dann das Noema mit dem fortgefahren werden sollte. Ziel ist eine repräsentative Menge der Daten zu sammeln um eine Klassifikation sinnvoll ausführen zu können. Trotz alle dem war die Beschränkung der Daten auch ausschlaggebend. Festgelegt wurde sich dann auf den Zeitraum eines Jahres und wegen der Zeit in der die Daten gesammelt wurden (Ende 2023) war das Jahr 2022 passend für die Aufgabe.

Das Projekt konnte nun in mehrere Schritte unterteilt werden: die Daten sammeln, die Daten aufbereiten, den Klassifikator programmieren, Testen und verbessern und die Arbeit zu dokumentieren.

Bei einer Besprechung mit dem Betreuer des Projekts Herrn Prof. Dr. Baumann das Projekt besprochen wurde wurde klar, dass eine Klassifizierung der Daten nicht das