

# H2D - SUCHE NACH UNTERNEHMEN MITTELS TEXT MINING UND DURCH IDENTIFIZIERUNG REGIONALER MERKMALE

Yuri Campbell, Karl Trela, Frank Pothen, Laura Brock (Wissenschaftliche Mitarbeiter\*innen)  
 Miriam Armin, Satiyabooshan Murugaboopathy (Wissenschaftliche/ Studentische Hilfskräfte)

## Text Mining Firmen in der Wasserstoffwirtschaft

Die Etablierung einer grünen Wasserstoffwirtschaft gilt als entscheidender Schritt für das Erreichen der Kohlenstoffneutralität bis 2050. Elektrolyseure, die Wasser in Wasserstoff und Sauerstoff aufspalten, sind Bestandteile einer gut funktionierenden Wasserstoffwirtschaft. In mehreren Schritten wurden daher regionale Faktoren identifiziert, die einen Einfluss auf das Wachstum der Wasserstoffwirtschaft haben. Zunächst wurde dazu eine erste Suche nach Herstellern von Wasserstoff-Elektrolyseuren und Unternehmen mit Bezug zu Wasserstoff im Allgemeinen durchgeführt. Aus den Ergebnissen wurde ein Algorithmus abgeleitet, der 1,1 Millionen Unternehmen aus der Hoppenstedt-Firmendatenbank nach ihrer Wahrscheinlichkeit, Elektrolyseure zu produzieren, ordnet. Zur Erstellung des Algorithmus wurde zunächst eine Stichprobe von Firmen Webseiten mit der jeweils höchsten und niedrigsten Wahrscheinlichkeit, mit Wasserstoff im Allgemeinen zu tun zu haben, zusammengestellt. Anschließend wurden die niedriger eingestuften Unternehmen und die Ankerbeispiele genutzt, um ein Textklassifikationsmodell zu trainieren. Der trainierte Klassifikator bewertet potentielle Unternehmen und ermittelt die Wahrscheinlichkeit, dass das Unternehmen im Bereich der Wasserstoff-Elektrolyse agiert. Eine detaillierte Untersuchung der hoch gerankten Firmen kann präzise Ergebnisse liefern.

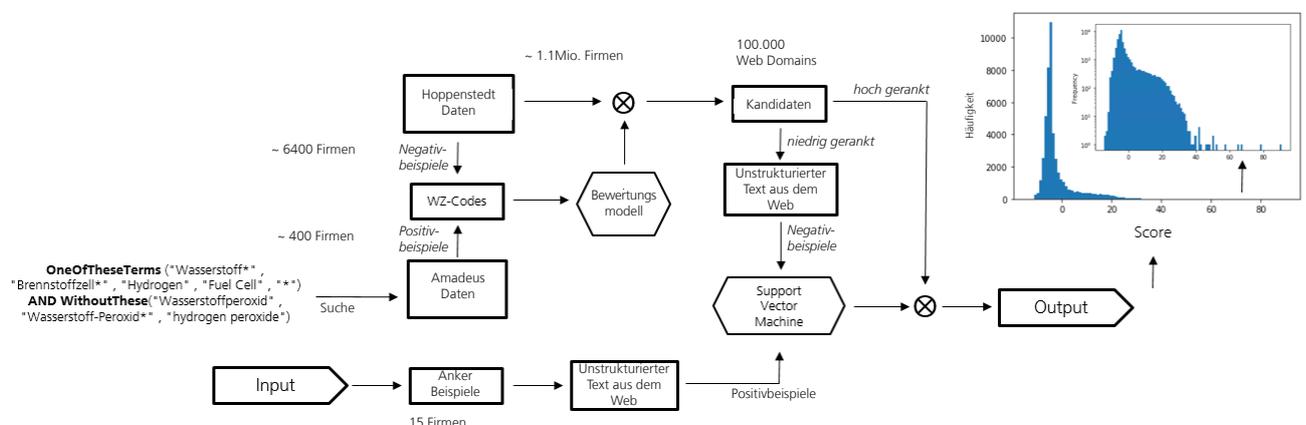


Abbildung 1: Hoppenstedt & Websuche Ansatz

## Wasserstoff und Elektrolyse in Deutschland

Auf Basis eines Analyse-Scores von mindestens 30 konnten 423 Unternehmen identifiziert werden, die im Bereich der Wasserstoff-Elektrolyse tätig sein könnten. Diese sind in der nebenstehenden Deutschlandkarte visualisiert. Potentielle Wasserstoff-Elektrolyseur-Firmen sind vor allem in Westdeutschland zu finden, wobei die Standorte eine Art Halbmond bilden, der sich von Hamburg über Köln, Frankfurt und Stuttgart bis nach München erstreckt. Ostdeutschland ist unterrepräsentiert, weist aber kleinere Cluster in Berlin und Sachsen auf. Eine Analyse der Anzahl der Firmen im Zeitverlauf könnte interessante Erkenntnisse für die Entwicklung der einzelnen Cluster ergeben.

### Standortwahl von Wasserstoff-Elektrolyseur-Firmen

Um regionale Merkmale zu identifizieren, die einen Einfluss auf die Anzahl der Unternehmen in der Branche haben, wurde die normierte Anzahl der Unternehmen, die im Hoppenstedt-Algorithmus einen Analyse-Score von über 50 erreichten, mit verschiedenen wirtschaftlichen Faktoren in den verschiedenen Bundesländern über einen Zehnjahreszeitraum von 2010-2019 korreliert. Abbildung 3 zeigt die Anzahl der Unternehmen pro Landkreis, unter Berücksichtigung der Bevölkerungsgröße, im Verhältnis zum BIP, zur Arbeitslosenquote, sowie zum Umsatz mit chemischen Produkten und zum Umsatz mit Metallerzeugung und -verarbeitung für das Jahr 2019. Die unteren Grafiken zeigen eine Regressionslinie, die die Poisson-Verteilung zur Schätzung eines verallgemeinerten linearen Modells verwendet. Der p-Wert für die Z-Statistik zeigt, dass das Vorhandensein von chemischer Industrie und einer starken Metallerzeugung und -verarbeitung signifikante Faktoren sind, die die Ausprägung einer grünen Wasserstoffindustrie stark beeinflussen. Unter Verwendung der Nagelkerke-Methode zur Bestimmung der Pseudo-R<sup>2</sup>-Werte für die jeweiligen Regressionen erhalten wir ein R<sup>2</sup> = 0,058 für die Regression für den Umsatz mit Chemikalien, ein R<sup>2</sup> = 0,034 für die Regression für den Umsatz mit Metallerzeugung und -verarbeitung, ein R<sup>2</sup> = 0,006 für die Regression mit der Arbeitslosenquote und ein R<sup>2</sup> = 0,045 für die Regression mit dem BIP. Dies deutet darauf hin, dass diese Faktoren zwar relevant, jedoch nicht die entscheidenden Determinanten für das Entstehen der grünen Wasserstoffindustrie sind.

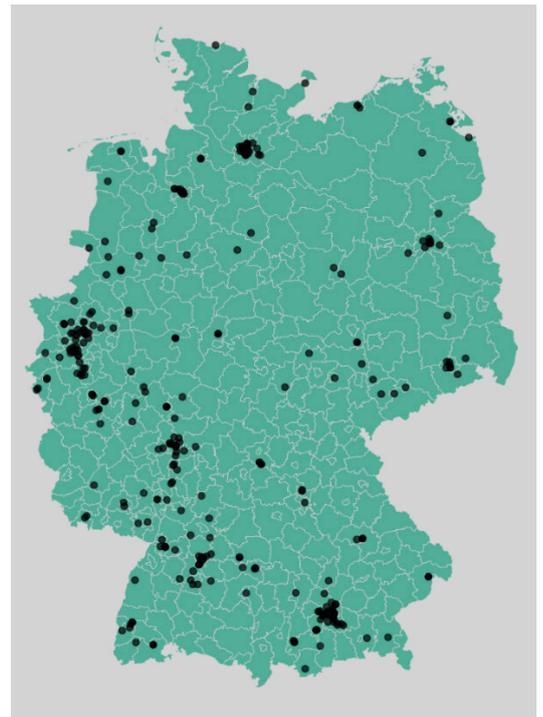


Abbildung 2: Verteilung der Elektrolyseur-Firmen

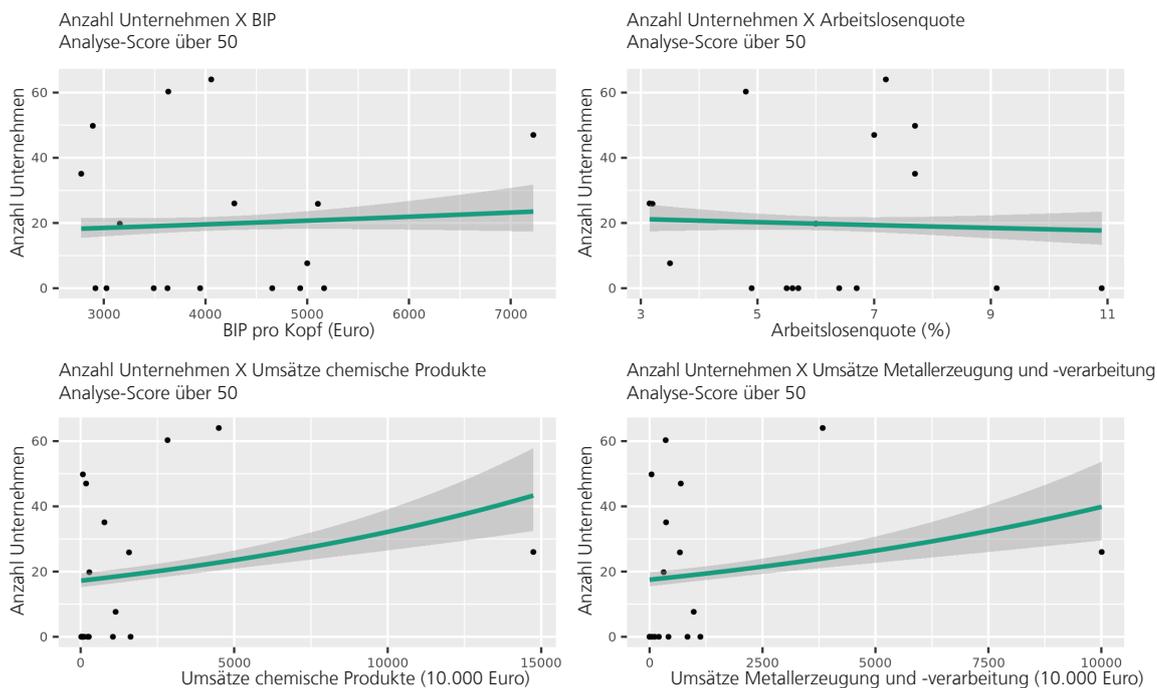


Abbildung 3: Poisson-Regression der Standortfaktoren