



Landeshauptstadt
München
IT-Referat
it@M

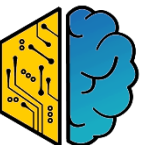
Mit Open Data zur KI

Merkmale und Trainingsdaten
aus offenen Daten generieren

Fabian Reinold – it@M InnovationLab



München.
Digital. Erleben.



- **Wie funktioniert KI & Maschinelles Lernen?**
- **Wie generiere ich aus Open Data Trainingsdaten?**
 - a) **Existierende Datensätze finden und kombinieren**
 - b) **Geo-Locations zu zusätzlichen Merkmalen transformieren**
 - c) **Bild-Datensätze mithilfe von Markierungen generieren**
- **Fragen und Diskussion**

Wie funktioniert KI & Maschinelles Lernen?

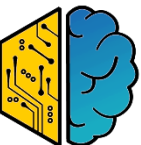


Landeshauptstadt
München
IT-Referat
it@M

Klassische Programmierung



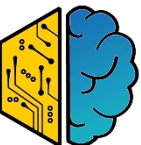
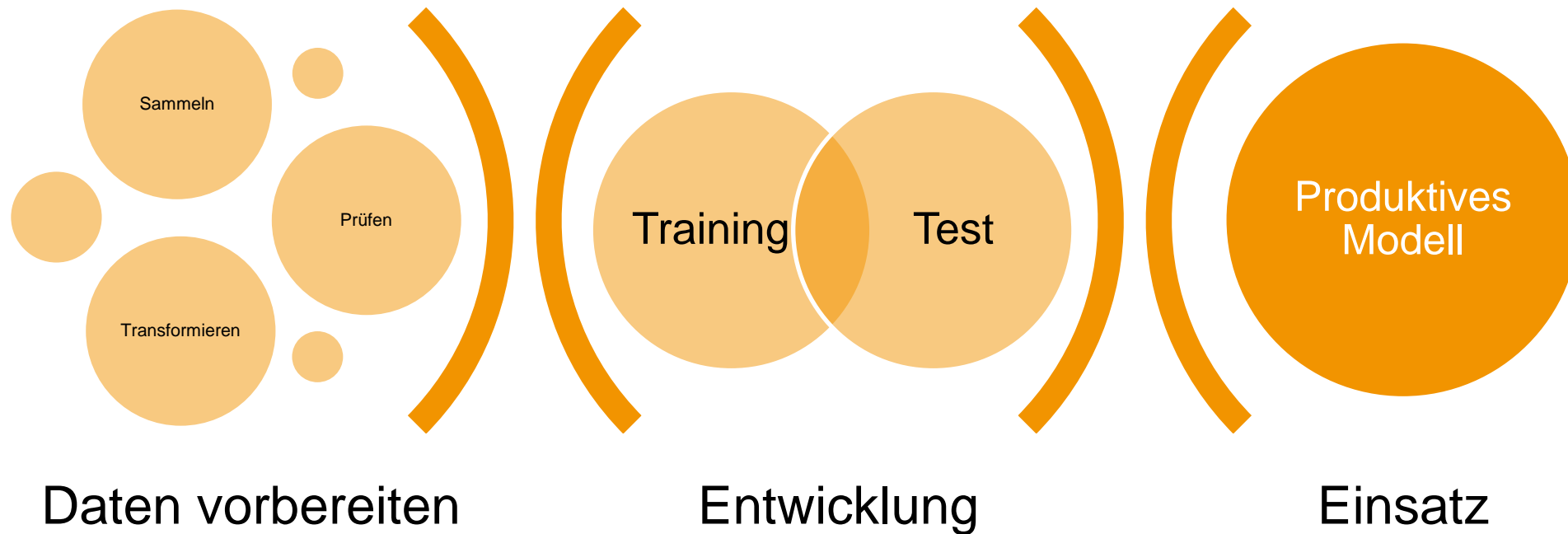
Maschinelles Lernen



Wie funktioniert KI & Maschinelles Lernen?



Landeshauptstadt
München
IT-Referat
it@M



Problemstellung beim Machine Learning:

Zum Training eines Modells benötigen wir nicht nur die Antworten (Zielwerte), sondern auch weitere Informationen (Merkmale)



Ein Beispiel:

Unser Ziel ist die **Vorhersage von Stickstoff-Werten** in München an LfU-Stationen.

Wir stellen uns dann die Frage: **Was hat Auswirkungen auf diesem Zielwert?**

Mögliche Antworten:

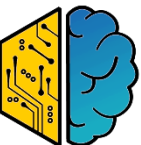
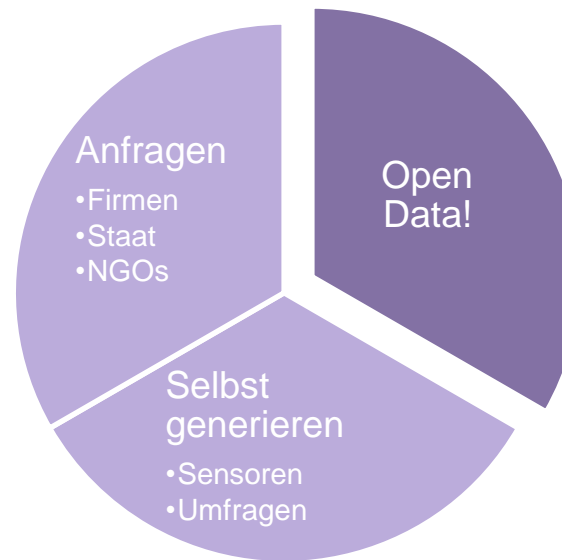
- Datum und Zeit
- Wetterbedingungen
- Emissionen aus Kraftverkehr
- Bebauung und Grünflächen in der Nähe

Wie funktioniert KI & Maschinelles Lernen?



Landeshauptstadt
München
IT-Referat
it@M

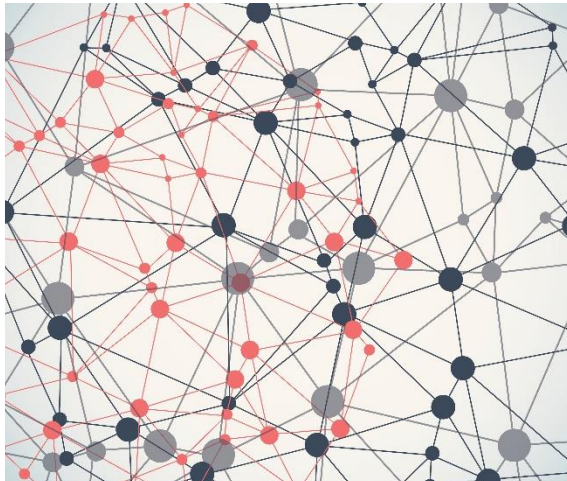
**Haben wir potentielle Merkmale identifiziert, gilt es diese zu beschaffen.
Aber wo?**



Wie generiere ich aus Open Data Trainingsdaten?



Landeshauptstadt
München
IT-Referat
it@M



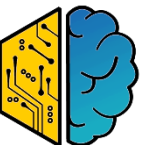
**Existierende Datensätze
finden und
transformieren**



**Geo-Locations zu
zusätzlichen Merkmalen
transformieren**



**Bild-Datensätze mithilfe
von Markierungen
generieren**

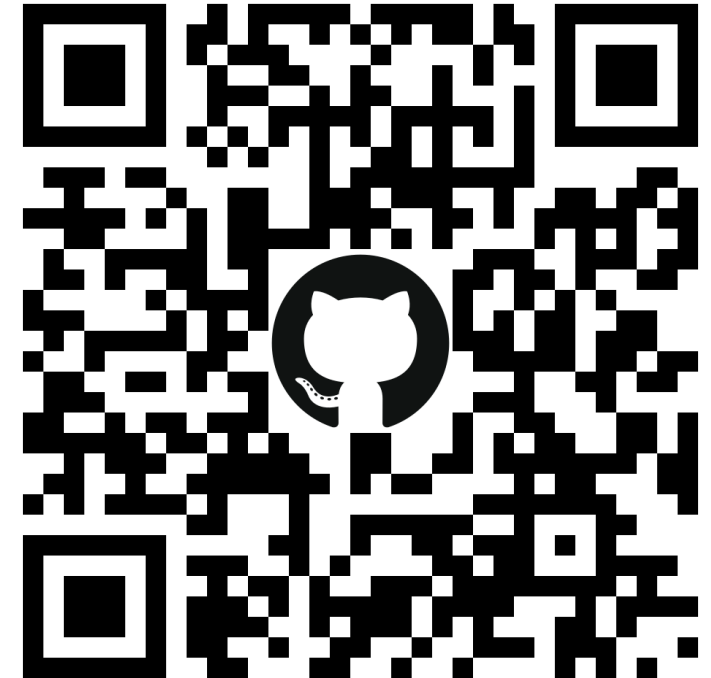


GitHub-Repository mit allen benutzten Daten und Skripten

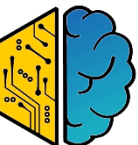


Landeshauptstadt
München
IT-Referat
it@M

<https://github.com/freinold/odd23-workshop>



München.
Digital. Erleben.

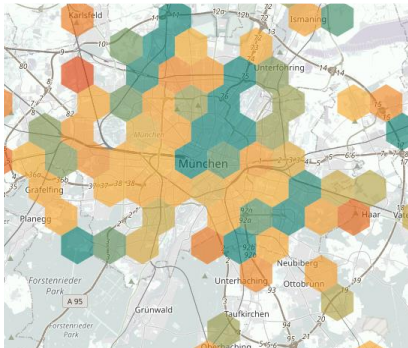


Beschaffung
von Feinstaub-
Daten beim LfU

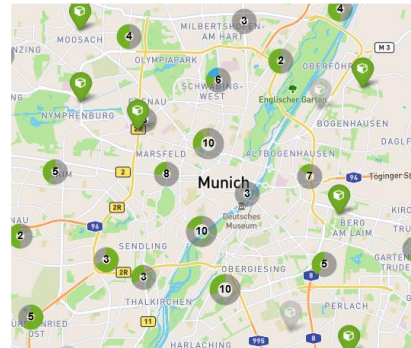
Zeitbezogene
Merkmale mit
muclearn
berechnen

Wetter-Daten
vom DWD
hinzufügen

Weitere Quellen für Wetter- & Umweltdaten:



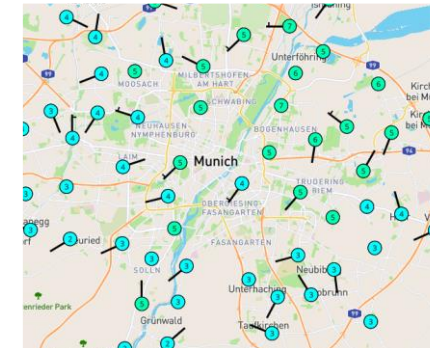
<https://sensor.community/de/>



<https://opensensemap.org/>



<https://www.awekas.at/>



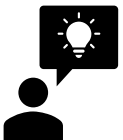
<https://www.wunderground.com/>



Szenario: Wir haben einen Datensatz mit Koordinaten in München und wollen die „Attraktivität“ der Punkte bestimmen.



Problem: Das Modell kann keine Karte öffnen und sich selbst „ein Bild davon machen“, wo die Punkte und zusammenhängende Orte liegen.



Lösung: Wir berechnen Quantisierungen der Position, mit denen das Modell arbeiten kann, z.B. kürzeste Distanz oder Anzahl im Umkreis

**Markierungen
von z.B.
Bäumen aus
OSM
extrahieren**

**Automatisches
Tagging
durchführen**

**auf Open Data
Luftbild
anzeigen**

**Trainingsdaten
für Objekt-
Erkennung auf
Luftbildern
erhalten**



Welche weiteren Szenarien für
Merkmalsgenerierung aus Open Data
könnt ihr euch vorstellen?





Herzlichen Dank für Ihre Aufmerksamkeit!



Fabian Reinold

fabian.reinold@muenchen.de



it@M InnovationLab

itm.innolab@muenchen.de

München.Digital.Erleben.



www.muenchen.digital



[@MuenchenDigital](https://twitter.com/MuenchenDigital)



www.youtube.com/muenchendigital



[@muenchen.digital](https://www.instagram.com/muenchen.digital)

