

HAW HAMBURG  
INFORMATIK MASTER

GRUNDPROJEKT

---

**TensorFlow Probability**

Evaluation der Bibliothek für  
probabilistische und statistische Analysen

---

*Bearbeiter:*

Tom Schöner (2182801)

*Betreuung:*

Prof. Dr. Olaf Zukunft

29. März 2019

# Inhaltsverzeichnis

<b>1</b>	<b>Abstract</b>	<b>3</b>
<b>2</b>	<b>Tensorflow Probability Komponenten</b>	<b>3</b>
2.1	Layers . . . . .	3
2.1.1	Layer 0: Tensorflow . . . . .	4
2.1.2	Layer 1: Statistical Building Blocks . . . . .	4
2.1.3	Layer 2: Model Building . . . . .	4
2.1.4	Layer 3: Probabilistic Inference . . . . .	4
<b>3</b>	<b>Benutzerfreundlichkeit</b>	<b>4</b>
3.1	Dokumentation . . . . .	4
<b>4</b>	<b>Integration in Tensorflow</b>	<b>4</b>
<b>5</b>	<b>Beispiel: Korrelation von Luftverschmutzung und Temperatur</b>	<b>4</b>
<b>6</b>	<b>Fazit</b>	<b>4</b>

**Abbildungsverzeichnis**

**Tabellenverzeichnis**

# 1 Abstract

Die auf Tensorflow basierende Bibliothek Tensorflow Probability<sup>1</sup> ermöglicht eine probabilistische Herangehensweise der Modellierung in Tensorflow. Mittels einer breiten Masse an vorhandenen Tools, wie statistischen Verteilungen, Sampling oder verschiedenster probabilistischer Keras Layer, können einfache bis hin zu komplexen Modellen erstellt werden. Berechnungen werden, wie man es aus Tensorflow gewohnt ist, durch *Dataflow Graphs*<sup>2</sup> abgebildet. Auf die verschiedenen Funktionsweisen und Schichten von Tensorflow Probability wird in Abschnitt 2 detaillierter eingegangen.

In dieser Evaluation soll die Bibliothek auf ihre Benutzerfreundlichkeit, was unter Anderem die Dokumentation einschließt, Handhabung beim Erstellen von statistischen Modellen und Integration in das Framework Tensorflow untersucht werden. Die Kategorie Maschinelles Lernen mit Hilfe von neuronalen Netzen ist hierbei als Schwerpunkt anzusehen, dabei werde ich auch gesondert auf die Integration für Keras eingehen.

## 2 Tensorflow Probability Komponenten

### 2.1 Layers

Die Struktur von Tensorflow Probability lässt sich, wie aus der Dokumentation zu entnehmen ist<sup>3</sup>, in die folgenden vier Schichten einteilen. Die Schichten bauen hierarchisch aufeinander auf, abstrahieren die unterliegenden Schichten aber nicht zwangsläufig. Möchte man beispielsweise durch *MCMC* in Schicht 3 Modelle erstellen, sollte man mit *Bijectors* aus Schicht 1 vertraut sein.

---

<sup>1</sup><https://www.tensorflow.org/probability>

<sup>2</sup><https://www.tensorflow.org/guide/graphs>

<sup>3</sup><https://www.tensorflow.org/probability/overview>

### **2.1.1 Layer 0: Tensorflow**

### **2.1.2 Layer 1: Statistical Building Blocks**

Verteilungen / Bijectors

### **2.1.3 Layer 2: Model Building**

Edward2 / Probabilistic Layers with Keras / Trainable Distributions

### **2.1.4 Layer 3: Probabilistic Inference**

MCMC / VI / Optimizers

## **3 Benutzerfreundlichkeit**

### **3.1 Dokumentation**

## **4 Integration in Tensorflow**

## **5 Beispiel: Korrelation von Luftverschmutzung und Temperatur**

Das Jupyter Notebook ist unter [https://github.com/tom-schoener/ml-probability/blob/master/tfp-evaluation/notebooks/air\\_quality.ipynb](https://github.com/tom-schoener/ml-probability/blob/master/tfp-evaluation/notebooks/air_quality.ipynb) einsehbar.

## **6 Fazit**