

HAW HAMBURG
INFORMATIK MASTER

GRUNDPROJEKT

TensorFlow Probability

Evaluation der Bibliothek für
probabilistische und statistische Analysen

Bearbeiter:

Tom Schöner (2182801)

Betreuung:

Prof. Dr. Olaf Zukunft

3. April 2019

Inhaltsverzeichnis

1	Abstract	3
2	Tensorflow Probability Komponenten	3
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
3	Pragmatik und Semantik	4
4	Integration in Tensorflow	4
5	Beispiel: Korrelation von Luftverschmutzung und Temperatur	4
6	Fazit	4

Abbildungsverzeichnis

Tabellenverzeichnis

1 Abstract

Die auf Tensorflow basierende Bibliothek Tensorflow Probability¹ erweitert das Framework um eine probabilistische Komponente. Mittels einer breiten Masse an bereitgestellten Tools, wie statistischen Verteilungen, Sampling oder verschiedenster probabilistischer Schichten für neuronale Netze, können einfache bis hin zu komplexen Modellen erstellt werden. Berechnungen werden, wie man es aus Tensorflow gewohnt ist, durch *Dataflow Graphs*² abgebildet. Auf die verschiedenen Funktionsweisen und Schichten von Tensorflow Probability wird in Abschnitt 2 detaillierter eingegangen.

In dieser Evaluation soll die Bibliothek auf ihre Semantik und Pragmatik, Effektivität beim Erstellen von statistischen Modellen und Integration in das Framework Tensorflow untersucht werden. Das maschinelle Lernen mit Hilfe von neuronalen Netzen und deren Abstraktion durch Keras ist hierbei als Schwerpunkt anzusehen.

2 Tensorflow Probability Komponenten

2.1 Layers

Die Struktur von Tensorflow Probability lässt sich, wie aus der Dokumentation zu entnehmen ist³, 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 Parameter seines probabilistischen Modells mittels Sampling ermitteln, sollten *Bijectors* aus Schicht 1 kein Fremdwort sein.

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

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

³<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 Pragmatik und Semantik

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 einsehbar.

6 Fazit