Stochastic Machine Learning
Chapter 02 - Universal Approximation theorems

Thorsten Schmidt

Abteilung für Mathematische Stochastik

www.stochastik.uni-freiburg.de
thorsten.schmidt@stochastik.uni-freiburg.de

SS 2024

Universal approximation theorems

After having finished the study of Hornik (1991) and Döhler and Rüschendorf (2001), we now look at the following literature:
Literature:

- Liu et al. (2024)
- Kratsios (2021)

Both papers are shortly viewed and disucussed.

Kolmogorov-Arnold Networks

The universal approximation property

We revisited the paper Kratsios (2021) who introduced the notion of an **architecture**.
We work on a function space $\mathcal{X}$. On this function space we have a set of functions $\mathscr{F}$.
Then there is a partial function (which simply maps from a subset to the output space)

$$\circlearrowleft : \cup_{i \geq 1} \mathscr{F}^i \to \mathcal{X}$$

such that it is not trivial: there exists on $f \in \mathcal{X}$ and $f_i \in \mathscr{F}$, such that

$$f = \circlearrowleft \left((f_j)_{j=1}^{i} \right)$$

Under a composition of the space and suitable approximation properties on the subspaces, the universal approximation property holds. Also any universal approximator can be approximated by a transformed neural network.

Döhler, Sebastian and Ludger Rüschendorf (2001). „An approximation result for nets in functional estimation". In: **Statistics & probability letters** 52.4, pp. 373–380.

Hornik, Kurt (1991). „Approximation capabilities of multilayer feedforward networks". In: **Neural networks** 4.2, pp. 251–257.

Kratsios, Anastasis (2021). „The universal approximation property: characterization, construction, representation, and existence". In: **Annals of Mathematics and Artificial Intelligence** 89.5, pp. 435–469.

Liu, Ziming et al. (2024). „Kan: Kolmogorov-arnold networks". In: **arXiv preprint arXiv:2404.19756**.