The NN theorem!

Universal Approximation theorem

Let

- $\varphi(\cdot)$ be a $\begin{cases} \text{nonconstant,} \\ \text{bounded,} \\ \text{monotonically-increasing,} \\ \text{continuous} \end{cases}$ function on $\mathbb{R}$.
- $I_p = [0,1]^p$.
- $C(I_p)$ be the space of continuous functions on I_p

The NN theorem!

Universal Approximation theorem

Let

- $\varphi(\cdot)$ be a $\begin{cases} \text{nonconstant,} \\ \text{bounded,} \\ \text{monotonically-increasing,} \\ \text{continuous} \end{cases}$ function on $\mathbb{R}$.
- $I_p = [0,1]^p$.
- $C(I_p)$ be the space of continuous functions on I_p

Then, $\forall f \in C(I_p)$, $\forall \varepsilon > 0$,

The NN theorem!

Universal Approximation theorem

Let

- $\varphi(\cdot)$ be a $\begin{cases} \text{nonconstant,} \\ \text{bounded,} \\ \text{monotonically-increasing,} \\ \text{continuous} \end{cases}$ function on $\mathbb{R}$.
- $I_p = [0,1]^p$.
- $C(I_p)$ be the space of continuous functions on I_p

Then, $\forall f \in C(I_p)$, $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\alpha_i \in \mathbb{R}^{p+1}, \beta_i \in \mathbb{R}$ $(i \in [1, N])$,

such that the function $F(x) = \sum_{i=1}^{N} \beta_i \varphi\left(\alpha_i^T x\right)$

is an approximate realization of the function f:$\forall x \in I_p,\ |F(x) - f(x)| < \varepsilon$.

The NN theorem!

Universal Approximation theorem

Wait. . . what?

The NN theorem!

Universal Approximation theorem

Let's say that again:

[Under reasonable assumptions and definitions]

$$F(x) = \sum_{i=1}^{N} \beta_i \varphi\left(\alpha_i^T x\right)$$

$$\forall x \in I_p,\ |F(x) - f(x)| < \varepsilon$$

The NN theorem!

Universal Approximation theorem

Let's say that again:
[Under reasonable assumptions and definitions]

$$F(x) = \sum_{i=1}^{N} \beta_i \varphi\left(\alpha_i^T x\right)$$

$$\forall x \in I_p,\ |F(x) - f(x)| < \varepsilon$$

In other words, functions of the form $F(x)$ are dense in $C(I_p)$.

The NN theorem!

Universal Approximation theorem

Duh. . . ?

The NN theorem!

Universal Approximation theorem

$$F(x) = \sum_{i=1}^{N} \beta_i \varphi\left(\alpha_i^T x\right)$$

Given a big enough budget on N, one can approximate any function f to any arbitrary precision ε using functions of the form $F(x)$!

The NN theorem!

Universal Approximation theorem

$$F(x) = \sum_{i=1}^{N} \beta_i \varphi \left(\alpha_i^T x \right)$$

Given a big enough budget on N, one can approximate any function f to any arbitrary precision ε using functions of the form $F(x)$!

Consequence:
With enough neurons, a single layer, feed-forward ANN can approximate any function to any precision. It is a Universal Approximator.

The NN theorem!

Universal Approximation theorem

$$F(x) = \sum_{i=1}^{N} \beta_i \varphi \left(\alpha_i^T x \right)$$

Given a big enough budget on N, one can approximate any function f to any arbitrary precision ε using functions of the form $F(x)$!

Consequence:
With enough neurons, a single layer, feed-forward ANN can approximate any function to any precision. It is a Universal Approximator.

Warning:
"with enough neurons" $\rightarrow N$ might be so large, the network might not be learnable in practice!