

Närvivõrkude lähendusteooria ja Barroni ruumid

Juhendajad: Rainis Haller ja Andre Ostrak

Milliseid funktsioone saab närvivõrkudega lähendada ja kui suurt võrku on soovitud täpsuse saavutamiseks vaja? Need küsimused kuuluvad masinõppe matemaatilistesse alustesse ning seovad klassikalise lähendusteooria funktsionaalanalüüsi ja tõenäosusteooriaga.

Projekt keskendub ühe peidetud kihiga närvivõrkudele

$$N_m(x) = \sum_{j=1}^m a_j \sigma(\langle w_j, x \rangle + b_j), \quad \sigma(t) = \max\{0, t\},$$

kus $x \in [0, 1]^d$, $w_j \in \mathbb{R}^d$ ja $a_j, b_j \in \mathbb{R}$. Funktsiooni σ nimetatakse ReLU-aktivatsioonifunktsiooniks ning m on peidetud kihi neuronite arv.

Töö eesmärk. Koostada iseseisvalt loetav referatiivne ülevaade närvivõrkude lähendamisevõimest, esitades valitud põhitulemuste täielikud tõestused ja selgitavad näited. Töö põhiosad on järgmised.

1. Universaalne lähendamine. Alustada ühe muutuja ositi lineaarsete funktsioonide esitamisest ReLU-võrkudena. Seejärel käsitleda teoreemi, mille kohaselt saab iga pidevat funktsiooni kuubil $[0, 1]^d$ selliste võrkudega ühtlaselt lähendada, lubades neuronite arvul kasvada.

2. Barroni ruum ja lähendamiskiirus. Tutvustada närvivõrkude integraalesitust ning selle kaudu määratud Barroni ruumi $\mathcal{B}$ ja normi. Tõestada lähendushinnang kujul

$$\|f - N_m\|_{L^2(\mu)} \leq \frac{C\|f\|_{\mathcal{B}}}{\sqrt{m}}, \quad f \in \mathcal{B},$$

kus μ on tõenäosusmõõt kuubil ning sobivate parameetritega võrk N_m leidub iga $m \geq 1$ korral. Selgitada juhusliku valiku ja kumerate kombinatsioonide osa tõestuses. Kasutatava Barroni normi korral saab konstandi C valida mõõtmest sõltumatult; funktsiooni norm ise võib mõõtmega kasvada.

3. Näited ja tulemuste tähendus. Arvutada lihtsate funktsioonide Barroni normi hinnanguid ning võrrelda lähendaja olemasolu ja lähendamiskiirust. Selgitada eraldi, millised küsimused puudutavad võrgu parameetrite leidmist ja andmetest õppimist.

Oodatav tulemus: terviklik tõestustele tuginev ülevaade, mis annab aluse edasiseks tööks närvivõrkudega seotud funktsiooniruumide ja nende geomeetria alal.

Eelteadmised: lineaaralgebra, matemaatiline analüüs ja tõenäosusteooria alused. Vajalikud funktsionaal-analüüsi ja mõõduteooria mõisted omandatakse töö käigus. Täpne maht kohandatakse tudengi ettevalmistusega.

Lähteallikad (valitud osad).

R. DeVore, B. Hanin ja G. Petrova, *Neural Network Approximation*, Acta Numerica **30** (2021), 327–444; arXiv:2012.14501.

Weinan E, Chao Ma ja Lei Wu, *The Barron Space and the Flow-induced Function Spaces for Neural Network Models*, eeskätt § 2; arXiv:1906.08039.

K. Spiliopoulos, R. Sowers ja J. Sirignano, *Mathematical Foundations of Deep Learning Models and Algorithms*, AMS, 2025; raamatu veebileht.