

Deep Learning Szeminárium

Rekurrens neurális hálózatok

Zombori Zsolt

Rényi Alfréd Matematikai Kutatóintézet

Szeretnénk figyelembe venni az egymást követő adatpontok közti összefüggéseket

- Működjön hosszú szekvenciákra
- Működjön változó hosszú szekvenciákra
- Időbeli eltolásra legyen invariáns

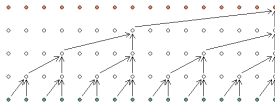
Idősorok modellezése 1D konvolúcióval

- Paramétermegosztás az időtengely mentén
- Csak sekély összefüggéseket tud modellezni a bemenetek között
- Dilated convolution: a hálózat mélységével exponenciálisan növekszik a visszatekintés mérete
- Előre rögzítjük az összefüggések terjedelmét
- Minden lépésnek a nyers adatokból kell dolgoznia

Non dilated Causal Convolutions



WaveNet Dilated Causal Convolutions



- A modell tartalmazza a múlt veszteséges összegzését egy $h(t)$ rejtett állapotban
- $h(t) = f(h(t-1), x(t), \theta)$
- Egyetlen f modell használható tetszőleges szekvencia hossz esetén
- Kimenet $y(t) = g(h(t))$ érkezik minden lépésben, vagy egyszer a végén.

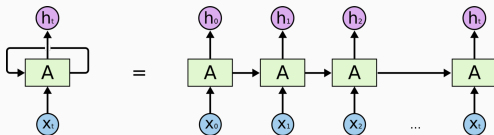
$h(t - 1)$ visszacsatolva:

- Turing teljes
- Előre- és visszaterjesztés alapvetően nem párhuzamosítható

$y(t - 1)$ visszacsatolva:

- Nem Turing teljes: a kimenetnek egyszerre kell a múltat tömörítenie és megfelelni a kimeneti elvárásnak.
- Tanítása párhuzamosítható. Valós kimenet helyett az elvárt kimenetet adjuk bemenetként (Teacher forcing)

Rekurrens hálózatok



- Hasonlít egy olyan előre csatolt hálózathoz, amiben bizonyos rétegek paraméterei meg vannak osztva
- Tanítás: Backpropagation Through Time (BPTT)
- Használhatósága nagyban függ attól, hogy milyen léptékű időbeli függőségek vannak az adatokban
- Hosszú függőségek esetén: eltűnő-felrobbanó gradiensek problémája

Long Short Term Memory Networks (LSTM)

- <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Cél: hosszútávú függőségek kezelése
- Ötlet: Fenntartunk egy csatornát, amin keresztül az információ (aktiváció és gradiens) könnyen áramlik
- Pont mind a reziduális hálózatokban (de az LSTM 20 évvel korábban megszületett)

Változó hosszúságú bemenet és kimenet szekvencia

- Gyakran más hosszúságú kimenetet várunk, mint bemenetet (pl nyelvek közti fordítás)
- Seq2seq: enkóder - dekóder architektúra

