

Meeting Notes - 2019.08.30

Før møtet

Faglig

- Tolkning av artikkelen
 - Analogi mellom nevralt nett og løsning av en differensialligning
 - * Stegene mellom lag i nettverket tolkes som tidssteg?
 - * Antar dette teoretisk kan gjøres kontinuerlig
 - * $f(u, t) = g_t(W_t y_{t-1} + b) \implies y_t = y_{t-1} + \Delta t g_t(W_t y_{t-1} + b)$, hvor f er uttrykket for den deriverte i en ODE.
 - * Vil lage en differensialligning som best mulig gir resultatene vi leter etter?
 - * Steglengdene i integratoren virker som en hyperparameter
 - * Er høyere ordens Runge Kutta metoder som å bruke større stider i et ResNet?
 - Invers lipschitz konstant, $\frac{1}{L}$, som fast steglengde for gradient descent?
 - * Dette er en ulik steglengde fra den som blir brukt i integratoren?
- Adjoint equation
 - Hva representerer p ?
 - Ressurser for å forstå adjoint equation og hamiltonian?
- Annen artikkel, Neural ODEs: <https://arxiv.org/pdf/1806.07366.pdf>

Kode

- Hvordan fungerer den romlige transformasjonen på dataen som vises i de genererte plottene?
 - Er det anvendt nettverket helt til outputnoden, så brukt at red-blue bakgrunnen representerer hypotesefunksjonen?

Brynjulf

- HBVP i Brynjulf sin kode?
 - Hamiltonian Boundary Value Problem
 - Hvordan brukes HBVP i `GradientCalc.m`? (Linje 23-43)

Matthias

- Backtrackingmetode med Lipschitz?

Generalisering av trening av den variable skrittlengden

- Vil vi trene en funksjon av t ?
- Skal jeg bruke backprop til å finne en derivert av loss mhp. $\vec{\Delta t}$?
- Skal hele nettverket trenes hver gang vi endrer t ?
- Meta-optimalisering

Under møtet

Variabelnavn fra kode

- DVfK = Derivative Vector field of K
- E = Runge Kutta koeffisient for metoden som brukes

Plan

- Foreslår følgende
 - Burde replikere kode i Tensorflow
 - Utlede uttrykk for å finne gradienten mhp. $\vec{\Delta t}$.
 - **Neste møte fredag 2019.09.06 kl. 14:30 norsk tid.**
- Brynjulf sender
 - Referanse på backtracking
 - Paper om gradient descent