



Vilniaus universitetas
Matematikos ir informatikos fakultetas
Informatikos katedra

Įvairūs gilieji neuroniniai tinklai

prof. dr. Olga Kurasova
Olga.Kurasova@mii.vu.lt

2018

Gilieji neuroniniai tinklai

- Nors pastaruoju metu **konvoliuciniai neuroniniai tinklai** yra vienas populiariausių giliųjų neuroninių tinklų tipas, tačiau **jie nėra vieninteliai** plačiai taikomi įvairiems uždaviniams spręsti.

Gilieji neuroniniai tinklai

Verta nepamiršti **šių neuroninių tinklų**:

- Deep Belief Network
- Deep Boltzman Machines
- **Deep Recurent Neural Network** (Long Short-Term Memory (LSTM) Networks)
- Deep Autoencoders

Data Science Ontology

click to expand or collapse

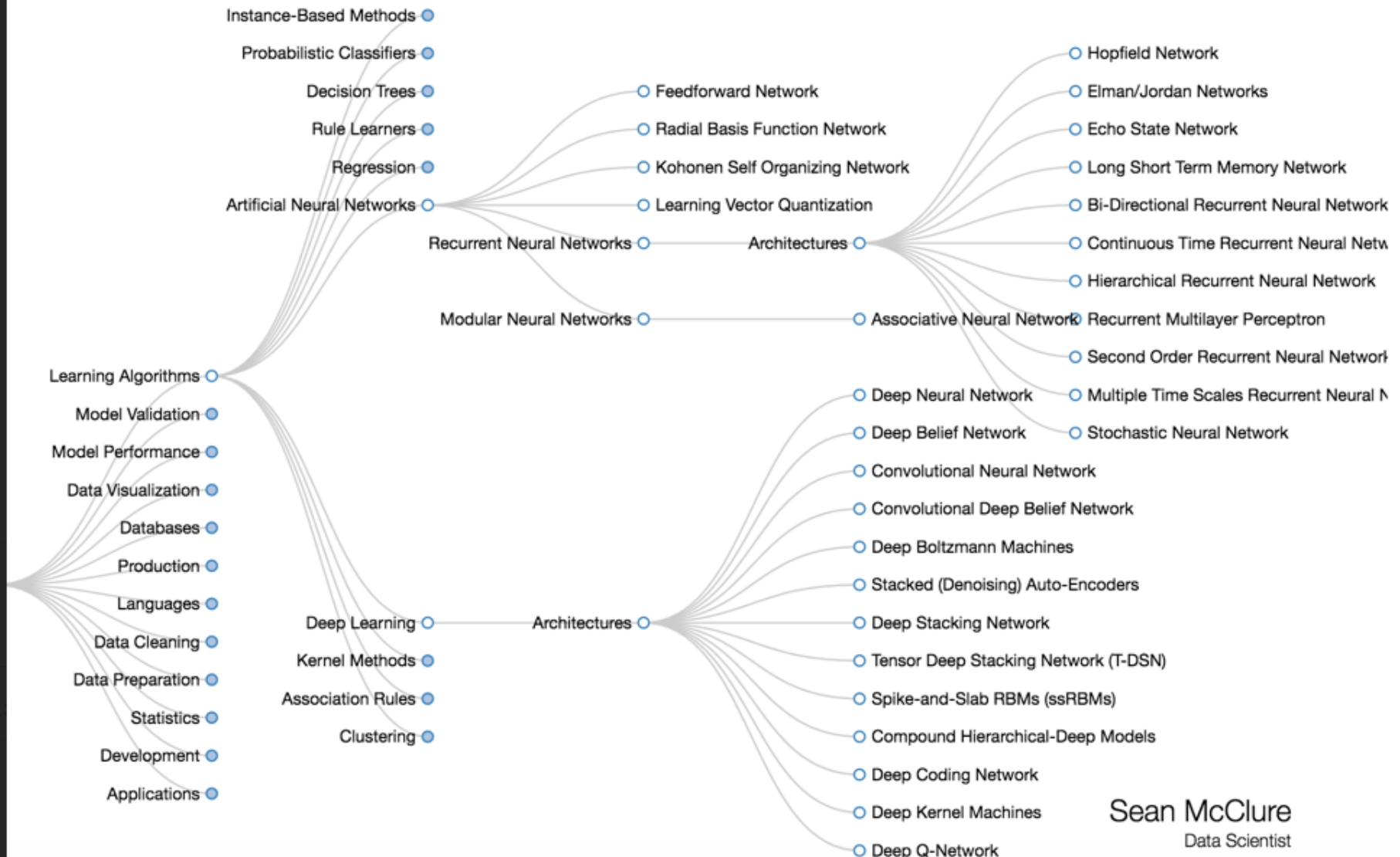
Tweet

Follow @WorldOfDataSci

Share

Close

click on terminal nodes to see wiki



Sean McClure
Data Scientist

Is <https://www.quora.com/What-are-the-types-of-deep-neural-networks-and-how-can-one-categorize-them-and-their-related-algorithms-as-either-shallow-or-deep>

Rekurentiniai neuroniniai tinklai

- **Rekurentiniai neuroniniai tinklai** (angl. *Recurrent Neural Networks, RNN*) – tai neuroniniai tinklai, kurių grafuose yra ciklų.
- Tai leidžia įgyvendinti **laike besikeičiančią veikseną**.
- RNN gali naudoti **jų vidinę atmintį** ir taip apdoroti į įvestis pateikiamas sekas.
- Tai leidžia spręsti tokius uždavinius, kaip **šnekos apdorojimas**.

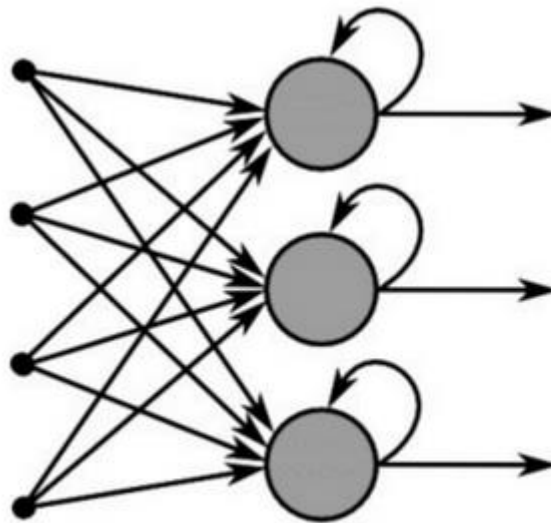
Rekurentiniai neuroniniai tinklai

- **Rekurentiniai neuroniniai tinklai** (*recurrent neural networks*, RNN) kuriame gana seniai (nuo 1980), tačiau pastaruoju metu, kaip ir visi giliojo mokymo neuroniniai tinklai, įgavo naują pagreitį.
- Tai tinklai, **turintys atmintį**, todėl gali puikiai tikti numatyti tai, kas bus ateityje.
- Įprastai jie turi trumpo laikotarpio atmintį (short-term memory), tačiau jie gali būti derinami su LSTM (long short-time memory).
- Jie naudojami **Apple Siri, Google Translate** ir kituose įrankiuose.

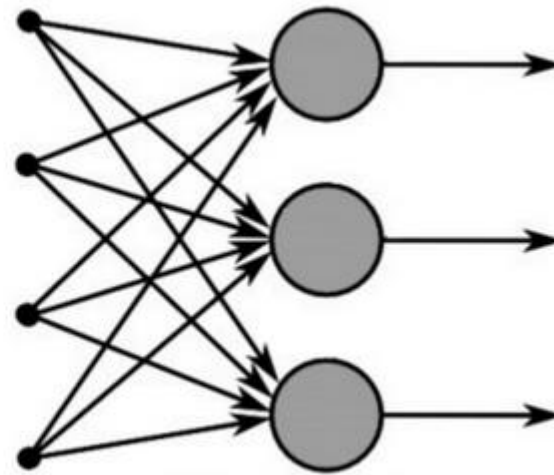
Rekurentinių neuroninių tinklų tipai (architektūros)

- **Ilgos trumpalaikės atminties tinklai** (angl. *long short-term memory (LSTM) networks*)
- Hopfieldo tinklai
- Elman/Jordan tinklai
- **Rekursyvūs neuroniniai tinklai** (angl. *recursive neural networks*)
- Rekurentiniai daugiasluoksniai perceptronai (angl. *recurrent multilayer perceptrons*)

RNN vs FFNN



Recurrent Neural Network

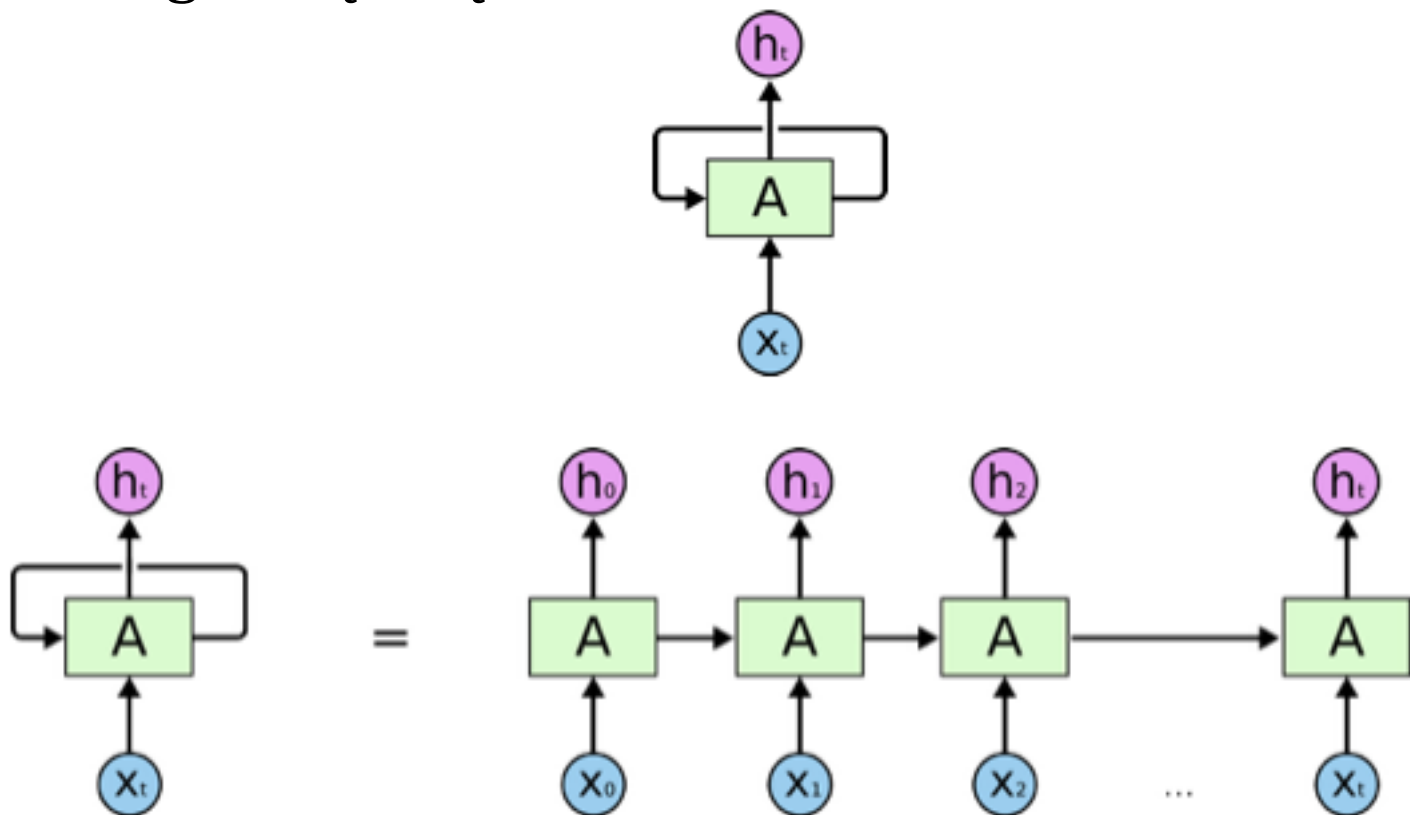


Feed-Forward Neural Network

<https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5>

Rekurentiniai neuroniniai tinklai

- **Rekurentiniai tinklai** turi ciklus, kurie leidžia informacijai būti perduotai iš vieno tinklo žingsnio į kitą.



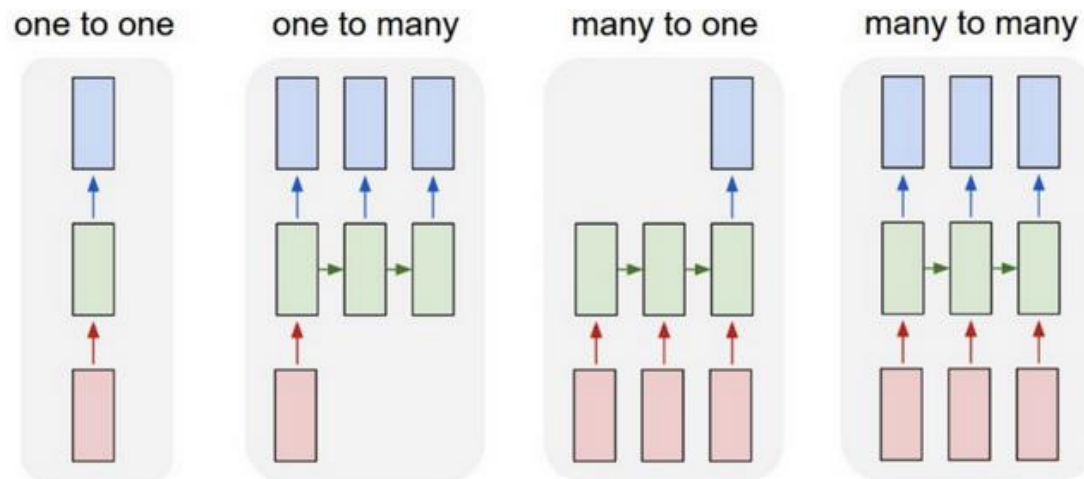
Pavyzdys

- „Imagine you have a normal feed-forward neural network and give it the word „neuron“ as an input and it processes the word **character by character**.“
- „At the time it reaches the character „r“, it has already **forgotten** about „n“, „e“ and „u“, which makes it almost impossible for this type of neural network to predict what character would come next.“

<https://towardsdatascience.com/recurrent-neural-networks-and-lstm-4b601dd822a5>

Rekurentiniai neuroniniai tinklai

- RNN mokymo metu pakeičiami svoriai ne tik **dabartinei įvesčiai**, bet ir **prieš tai buvusiai**.
- FFNN atvaizduoja **vieną įvestį į vieną išvestį**.
- RNN gali atvaizduoti vieną įvestį į daug išvesčių, daug į daug (vertimo atveju) arba daug į vieną (pvz., klasifikuojant balsą).



Ilgos trumpalaikės atminties (LSTM) tinklai

- Rekurentinių neuroninių tinklų praplėtimas yra **ilgos trumpalaikės atminties** (Long Short-Term Memory, LSTM) tinklai.
- Pavadinimas kilo nuo to, kad tai yra **trumpalaikės atminties, kuri gali tęstis ilgą periodą**, modelis.
- Modelis gerai tinka kai norima klasifikuoti, apdoroti ir prognozuoti **laiko eilutes** su nežinomo dydžio **vėlavimais** laike tarp svarbių įvykių.

Ilgos trumpalaikės atminties (LSTM) tinklai

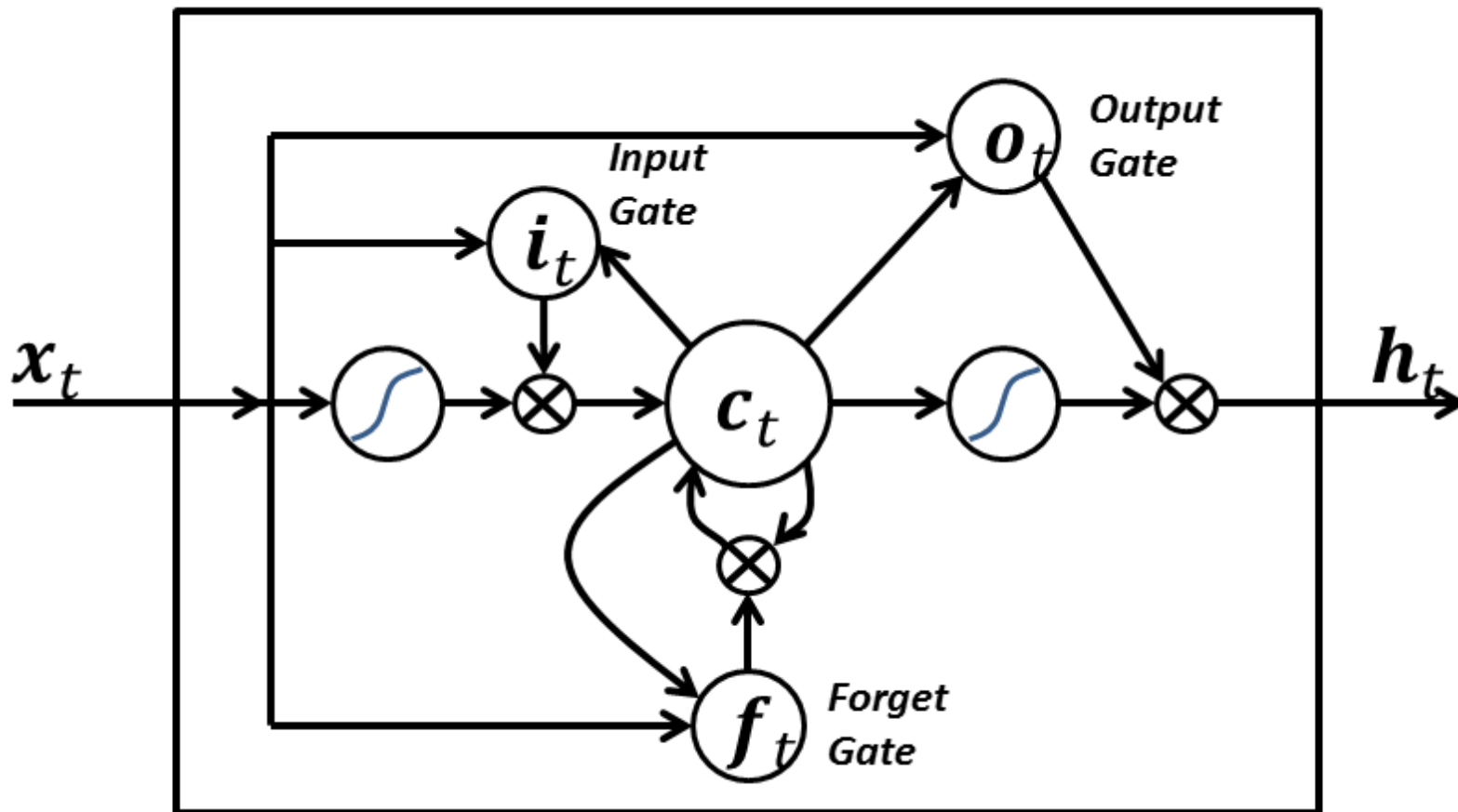
- Šių tinklų **pradžia** siekia **1997** metus.
- **2000** metais jie buvo **patobulinti**.
- Tokios IT srityje lyderiaujančios kompanijos, kaip **Google**, **Apple** ir **Microsoft** naudoja LSTM kaip pagrindinį komponentą daugelyje savo produktų.

Ilgos trumpalaikės atminties (LSTM) tinklai

Ivairūs taikymai:

- Robotų valdymas
- Laiko eilučių prognozavimas
- Šnekos atpažinimas
- Muzikos kūrimas
- Ranka rašomo teksto atpažinimas
- Žmogaus veiksmų atpažinimas
- Baltymų homologijos nustatymas
- Baltymų vietų prognozavimas

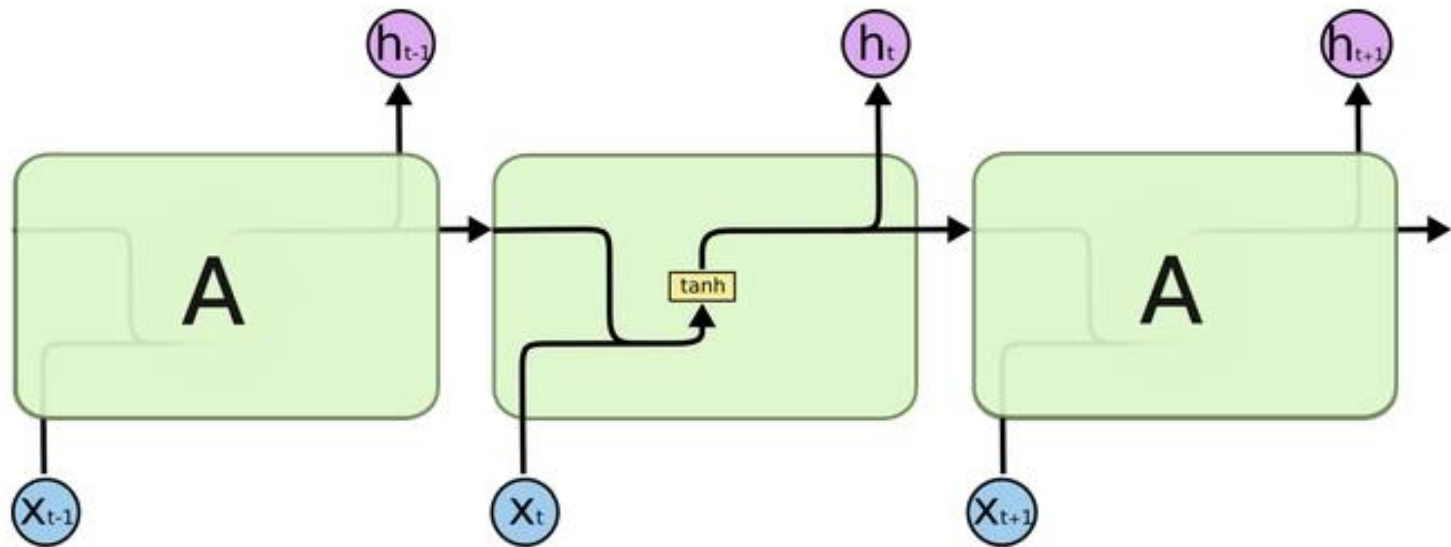
Ilgos trumpalaikės atminties (LSTM) tinklai



Long short-term memory (LSTM) block

Standard RNN vs LSTM

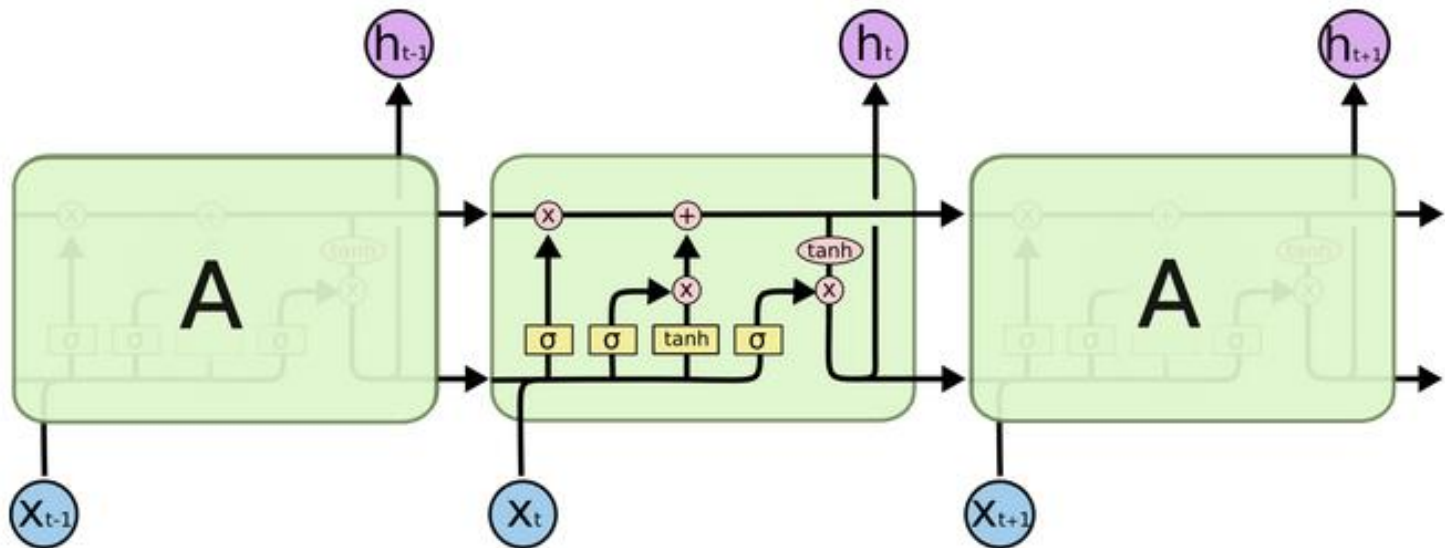
- „All RNNs have the form of a chain of repeating modules of neural network. In standard RNNs, this repeating module will have a **very simple structure**, such as a single *tanh* layer.“



The repeating module in a standard RNN contains a single layer.

Standard RNN vs LSTM

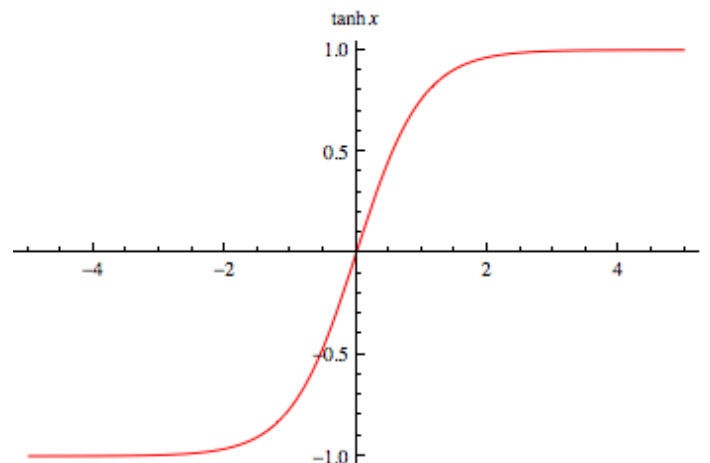
- „LSTMs also have this chain like structure, but the repeating module has a **different structure**. Instead of having a single neural network layer, there are **four**, interacting in a very special way.“



The repeating module in an LSTM contains four interacting layers.

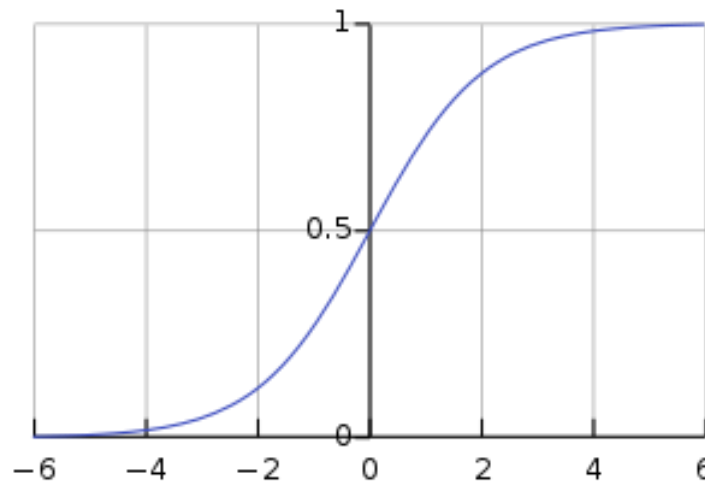
Kodėl \tanh ?

- Siekiant išvengti **gradiento nykimo problemos** (*vanishing gradient*), reikia naudoti funkciją, kurios antroji išvestinė gali išsilaikyti nelygia nuliui gana ilgą intervalą.
- **Hiperbolinis tangentas** \tanh atitinka šį reikalavimą.



Kodėl sigmoidinė funkcija?

- **Sigmoidinė funkcija** gali įgyti reikšmes 0, 1, kurios gali būti naudojamos „pamiršti“ arba „prisiminti“ informaciją.



Ilgos trumpalaikės atminties (LSTM) tinklai

- **Google** naudoja LSTM šnekai atpažinti telefone, išmaniajame asistente **Allo** (<https://allo.google.com>) ir **Google Translate**.
- **Apple** naudoja LSTM „**Quicktype**“ funkcijai įgyvendinti telefonuose iPhone.
- **Amazon** naudoja LSTM kaip **Amazon Alexa**.

Amazon Alexa

Alexa is Amazon's cloud-based voice service available on tens of millions of devices from Amazon and third-party device manufacturers. With Alexa, you can build natural voice experiences that offer customers a more intuitive way to interact with the technology they use every day. Our collection of tools, APIs, reference solutions, and documentation make it easy for anyone to build with Alexa.

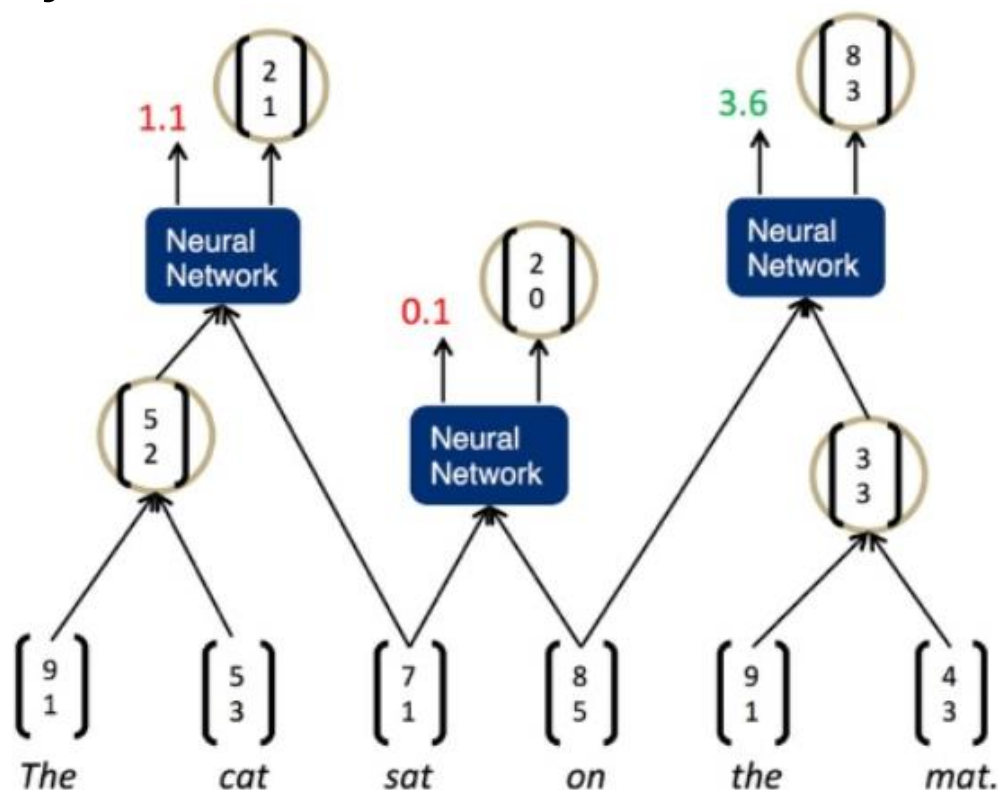
Iš <https://developer.amazon.com/alexa>

Rekursyvūs neuroniniai tinklai

- **Rekursyvūs neuroniniai tinklai** (angl. *recursive neural networks, RvNN*) – tai giliųjų neuroninių tinklų tipas, kuriuose rekursyviai dalijamasi svoriais.
- RvNN sėkmingai taikomas **natūralios šnekos apdorojime**, taip pat uždaviniuose, kuriose būtina apdoroti sekas.

Rekursyvūs neuroniniai tinklai

- Jie turi **medžio struktūrą**, kurioje kiekviename mazge yra neuroninis tinklas.



Iš <https://www.slideshare.net/jiessiecao/parsing-natural-scenes-and-natural-language-with-recursive-neural-networks>