

Boti un mākslīgais intelekts

atvijas Universitātes 76. konference
Datorlingvistikas sekcija

dr. Raivis Skadiņš, dr. Kaspars Balodis, Endijs Raikstiņš

1.03.2018



**LATVIJAS
UNIVERSITĀTE**
ANNO 1919

NACIONĀLAIS
ATTĪSTĪBAS
PLĀNS 2020



EIROPAS SAVIENĪBA
Eiropas Reģionālās
attīstības fonds



IEGULDĪJUMS TAVĀ NĀKOTNĒ

Šajā prezentācijā

kas ir boti

galvenie jēdzieni

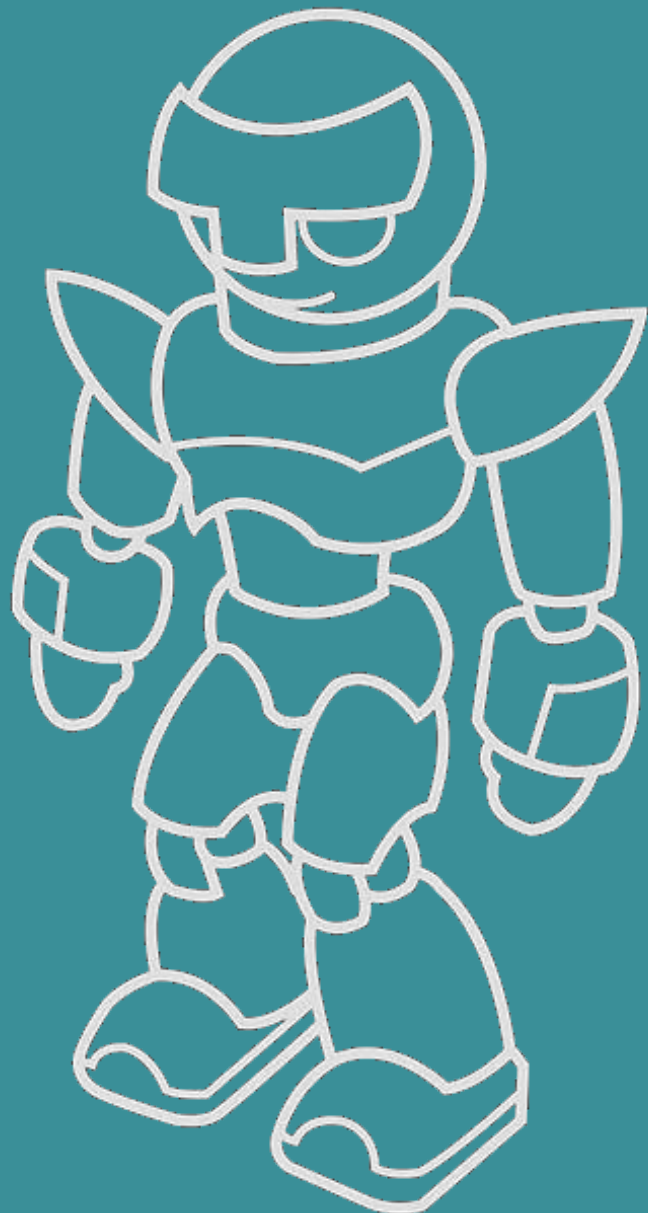
mākslīgais intelekts botos,
jeb nodoma noteikšana latviešu valodā

izaicinājumi

kas ir kas

mākslīgais intelekts

boti



daži piemēri

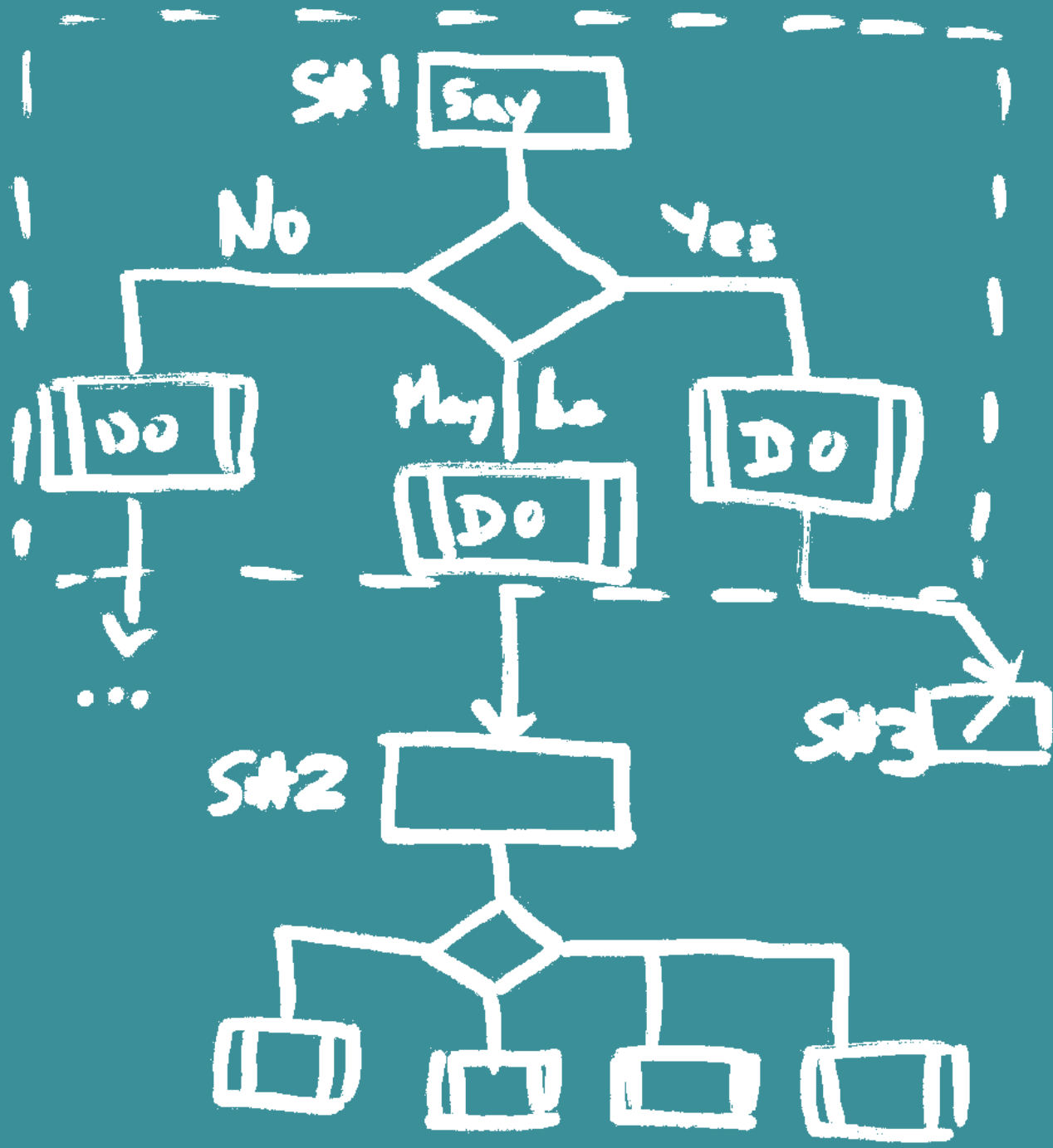
<https://www.messenger.com/t/VirtualAgentLaura>

<https://www.messenger.com/t/LattelecomAnete>

<https://www.messenger.com/t/ReizrekinaBots>

<https://www.messenger.com/t/LacisTeo>





galvenie komponenti

dialoga vadība (*dialog/conversation management/flow*)

nolūki (*intents*)

entītijas (*entities*)

darbības (*actions*)

nolūki

Ko lietotājs ar šo gribēja pateikt?

labdien/ sveiki/ čau /  / ...

welcome

jā/ protams/ kā citādi/ 😊/ jaa/ ...

yes

nē/ nopp/ atšujies/ ...

no

gribu pasūtīt picu/ vai jums ir pica?/ ...

order-pizza

līdz cikiem jūs strādājam

working-hours

vai varat piegādāt?

delivery

entītijas

sveiks, esmu Jānis Bērziņš

gribu divas picas

lai iet margarita šodien

ņemšu trīskāršo siera picu ar papildus sieru

uz vietas

vai strādājat līdz 10tiem

man lūdzu hawai uz puspieciem rītvakar

Jānis Bērziņš

2

margarita; 25.01.2018

3x-siera; ex-siera

on-place

10:00

hawai; 16:30;
26.01.2018

Un tas viss kopā arī ar nodomiem

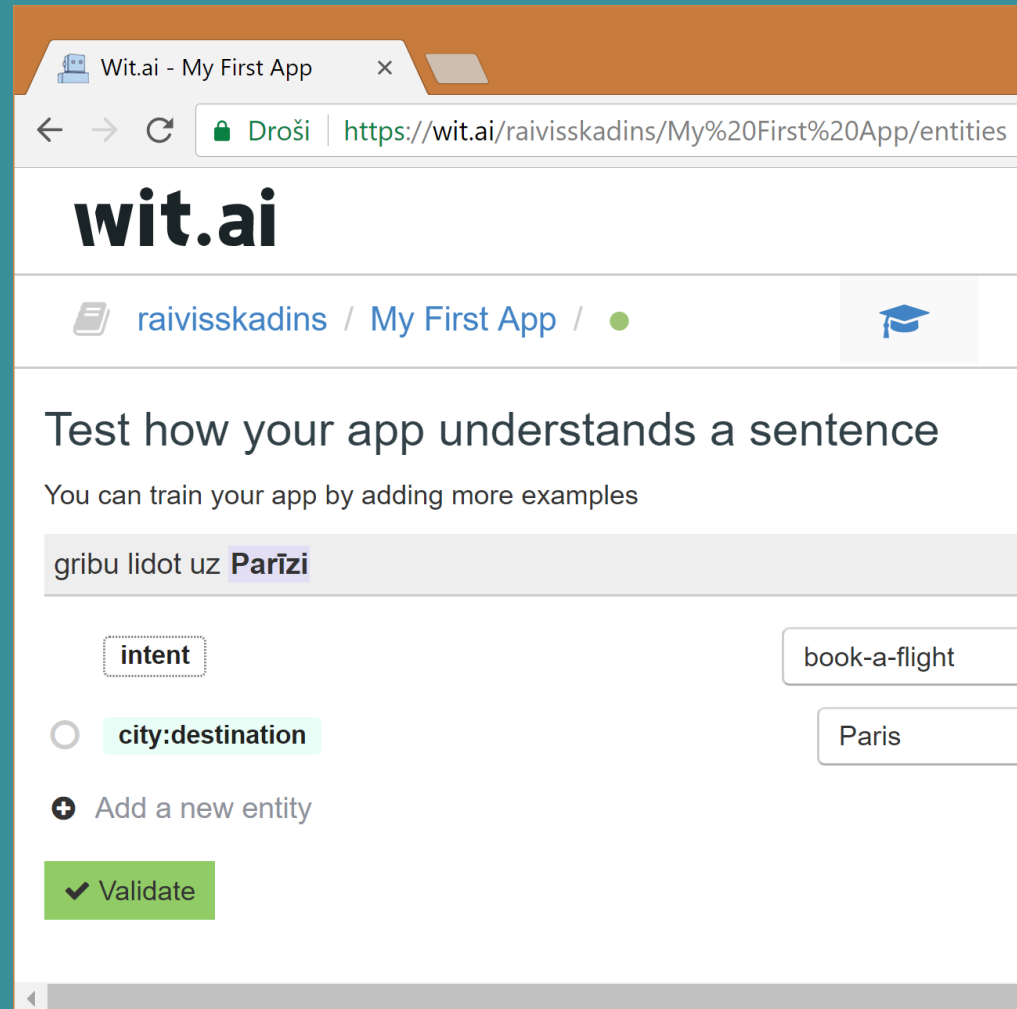
mašīnmācīšanās & AI

nolūku mācīšanās

entitāšu mācīšanās

dialoga plūsmu
vai pat visu dialogu
iēmācīšanās no piemēriem

end-to-end dialoga sistēmas



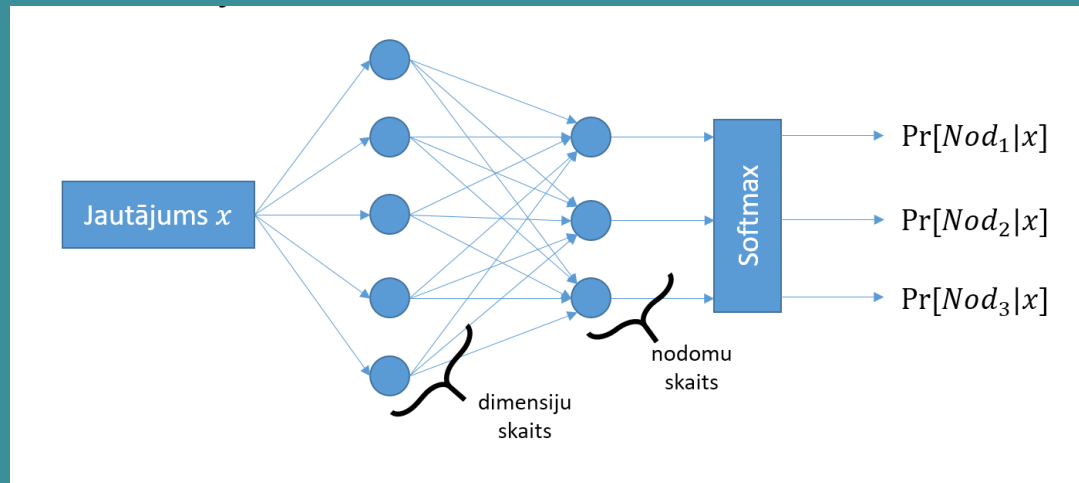
nolūka noteikšana

vektORIZĒŠANA

(onehot, fastText, MT, u.c.)

KLASIFICĒŠANA

(Naive Bayes, fastText, u.c.)



ĪPATNĪBAS:

locījumi, pieturzīmes, lielie/mazie burti, liekvārdi, drukas kļūdas u.c.

nolūka noteikšana

Metode	Reāli vērtēšanas dati	Salaboti vērtēšanas dati
wit.ai	36.44%	38.14%
OneHot vekt. + NaiveBayes	28.39%	32.20%
FastText vekt. un klasifikators	28.14%	30.59%
FastText vekt. un klasifikators + lemmatizācija + liekvārdu filtrs	32.29%	38.14%
FastText vekt. + ārējs klasifikators	23.73%	24.49%
FastText vekt. + ārējs klasifikators + lemmatizācija + liekvārdu filtrs	36.95%	41.86%
FastText vekt. + ārēja teikumu vekt. + ārējs klasifikators	38.22%	38.73%
FastText vekt. + ārēja teikumu vekt. + ārējs klasifikators + lemmatizācija + liekvārdu filtrs	38.73%	43.39%

izaicinājumi un turpmākie pētījumi

klasifikatori, citi neironu tīkli, piemēram, CNN, RNN

treniņdatu apjoms un kvalitāte

drukas kļūdas lietotāju ievadītajos tekstos

teksti, kas neatbilst nevienam paredzētajam *nodomam*

citas datu kopas un valodas



Projekts: “Neironu tīkli fleksīvo dabisko valodu apstrādei”
Projekta Nr.: 1.1.1.1/16/A/215



**LATVIJAS
UNIVERSITĀTE**
ANNO 1919

NACIONĀLAIS
ATTĪSTĪBAS
PLĀNS 2020



EIROPAS SAVIENĪBA
Eiropas Reģionālās
attīstības fonds



IEGULDĪJUMS TAVĀ NĀKOTNĒ