



Vilniaus miesto nusikaltimų prognozė

DARBĄ ATLIKO: TOMAŠ MONKEVIČ

DARBO VADOVAS: ASIST. DR. VYTAUTAS VALAITIS

DARBO RECENZENTAS: J. ASIST. LINAS PETKEVIČIUS

Panašūs tyrimai

- Autoįvykių nusikaltimų, kur nusikaltėlis pasišalino iš eismo įvykio vietos, prognozė naudojant „Twitter“ naudotojų pranešimus.
- San Francisko nusikaltimų tendencijų kitimo stebėjimas, nusikaltimų įtakojančių atributų nustatymas ir nusikaltimų prognozavimas.
- Amsterdamo nusikaltimų prognozė naudojant daugiasluoksnį perceptroną.
- Gvatemalos nusikaltimų prognozė naudojant ilgos trumpalaikės atminties modelį (LSTM).

Tikslas ir uždaviniai

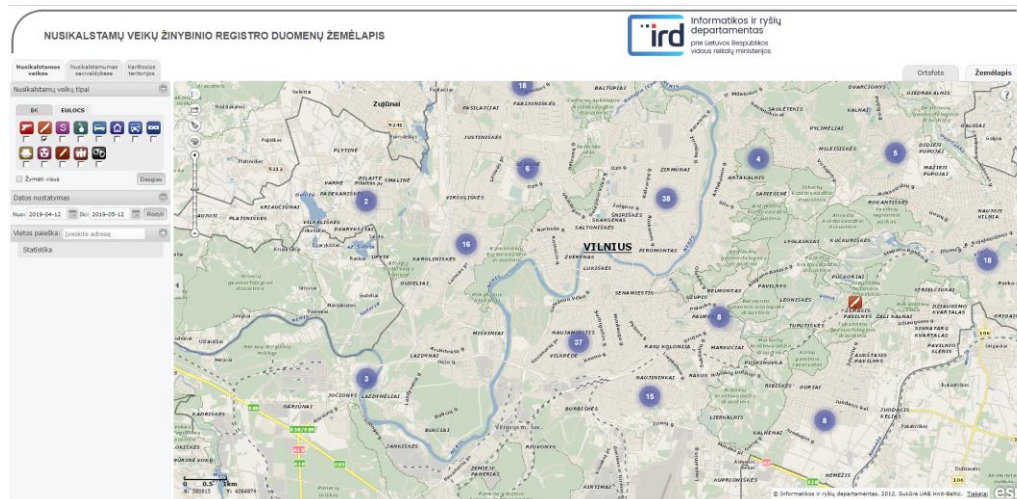
Darbo tikslas yra sukurti ir apmokyti ilgos trumpalaikės atminties (LSTM) modelį, kuris gebės prognozuoti nusikaltimų vietas kitai dienai Vilniuje.

Šiam tikslui pasiekti buvo iškelti trys uždaviniai:

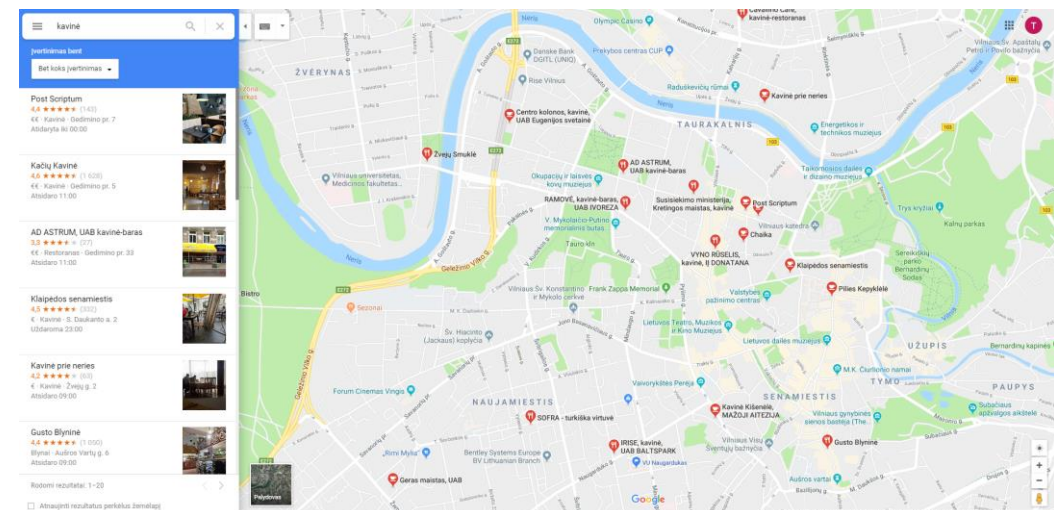
1. Surinkti reikalingus duomenis ir paruošti juos LSTM modeliui.
2. Pasirinkti LSTM modelio tikslumo vertinimo strategiją.
3. Sukurti ir apmokyti LSTM modelį.

Duomenų rinkimas (1)

- Vilniaus miesto nusikaltimų duomenys buvo gauti iš nusikalstamų veikų žinybinio registro duomenų žemėlapiu.
- Aplinkos duomenys buvo gauti iš „Google Places“ paslaugos.

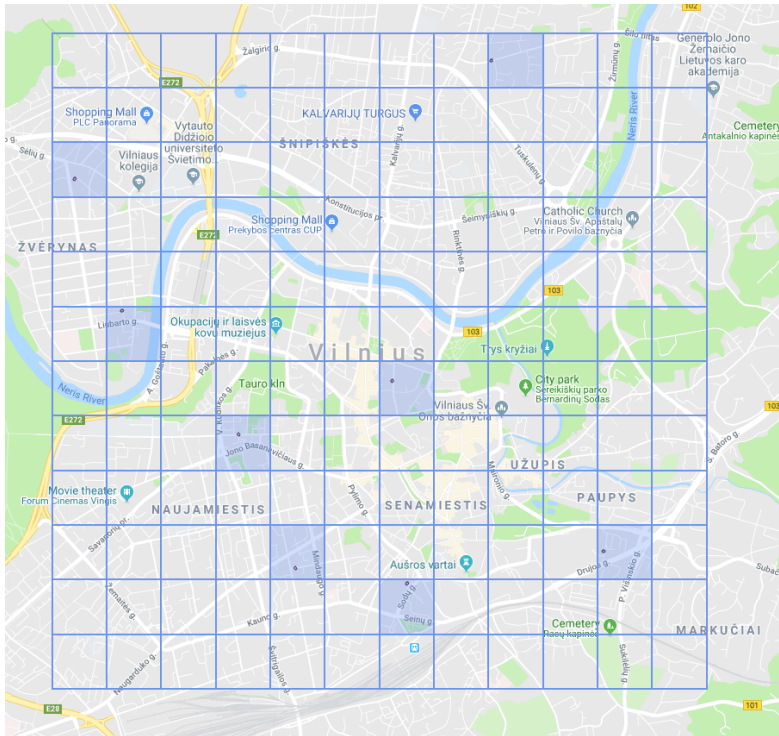


1 pav. Ekrano vaizdas iš nusikalstamų veikų žinybinio registro duomenų žemėlapiu

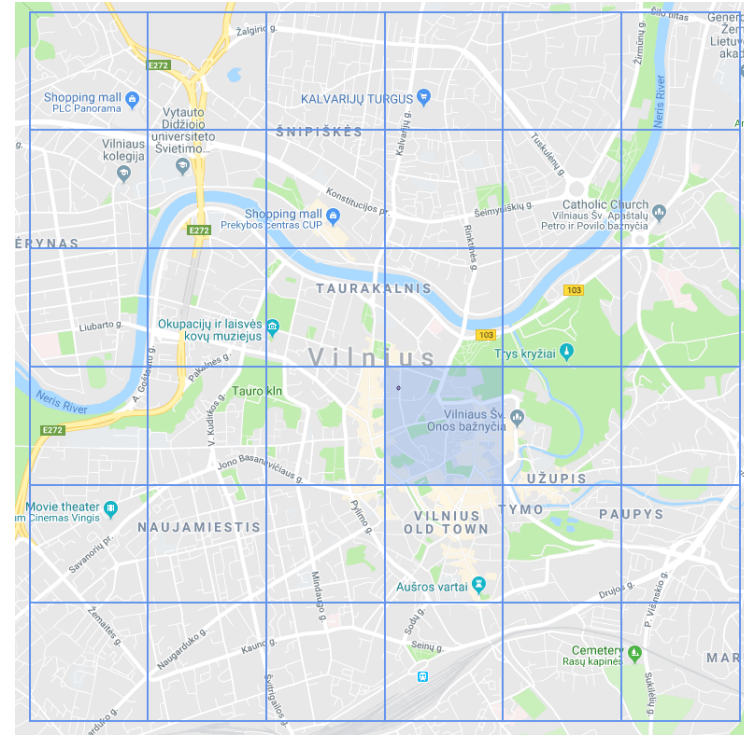


2 pav. Ekrano vaizdas iš Google žemėlapiu

Duomenų rinkimas (2)



(A) Tinklelio dydis 350x350 m²



(B) Tinklelio dydis 700x700 m²

3 pav. Sugrupuoti nusikaltimai į tinklėlį (bendras plotas 4.2x4.2 km²)

Modelio vertinimo strategija

Buvo naudojami keturi matai, kurie parodė modelio prognozės tikslumą:

- bendras tikslumas (*angl. accuracy*);

$$\frac{\sum_{i=1}^l \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l} \quad (1)$$

- prisiminimas (*angl. recall*);

$$\frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l (tp_i + fn_i)} \quad (2)$$

- pataikymo tikslumas (*angl. precision*);

$$\frac{\sum_{i=1}^l tp_i}{\sum_{i=1}^l (tp_i + fp_i)} \quad (3)$$

- F1 matas (*angl. F-measure*).

$$\frac{(\beta^2 + 1) * PataikymoTikslumas * Prisiminimas}{\beta^2 PataikymoTikslumas + Prisiminimas} \quad (4)$$

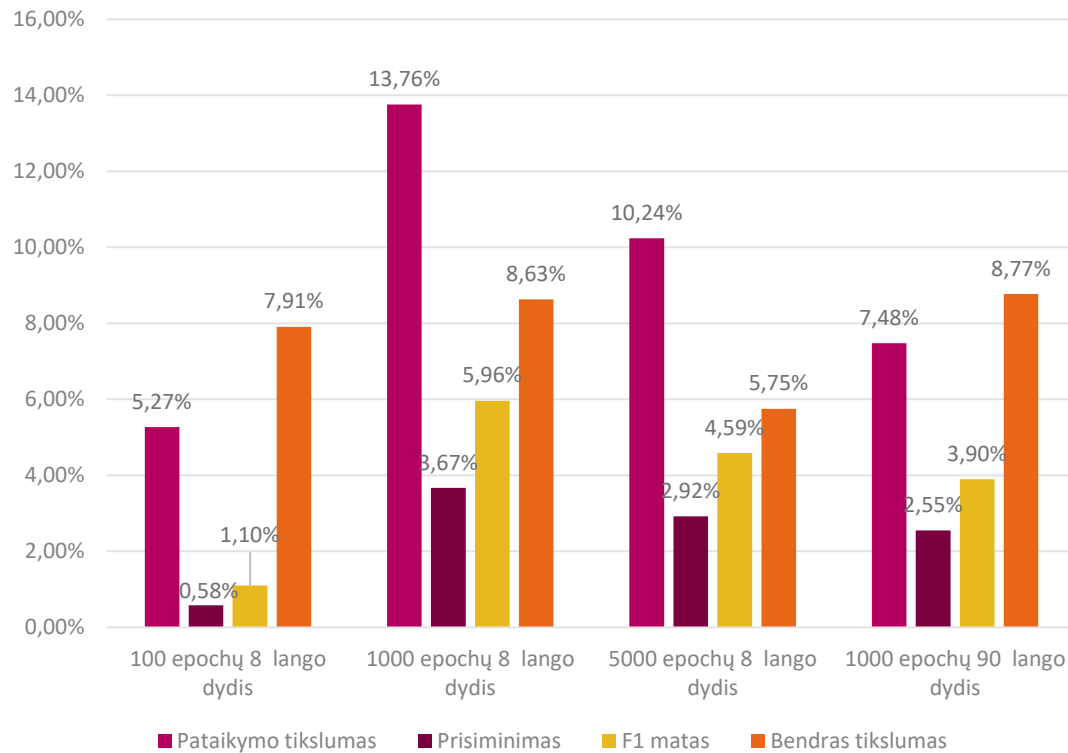
Modelio pritaikymas ir apmokymas (1)

1 lentelė. Šio darbo LSTM modelio ir kitų šaltinių prognozavimo rezultatai

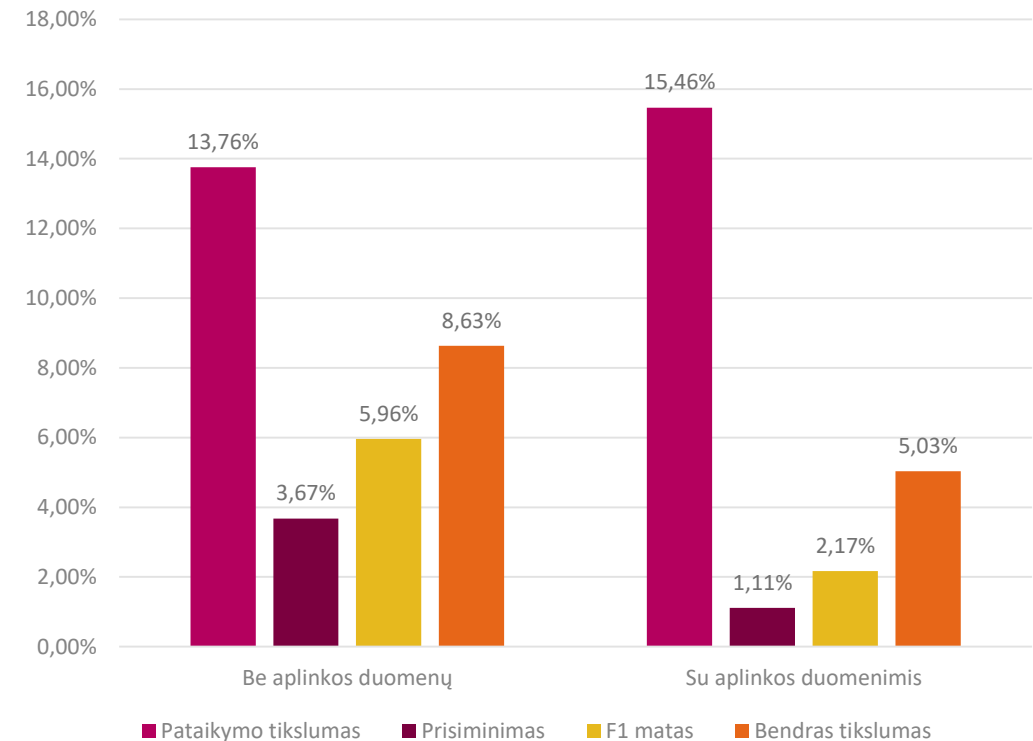
Prognozavimo modelis		Pataikymo tikslumas	Prisiminimas	F1 matas	Bendras tikslumas
*Daugiasluoksnis perceptronas su dvejų savaitių periodu		31,78%	53,19%	39,28%	-
*Daugiasluoksnis perceptronas su mėnesio periodu	Nakties	55,68%	70,18%	61,73%	-
	Dienos	58,52%	70,48%	63,64%	-
*Tarp grupių priklausomas binarinis LSTM modelis		60,71%	60,77%	60,73%	63,88%
Šio darbo LSTM modelis		5,27%	0,58%	1,10%	7,91%

* - kito šaltinio modelis

Modelio pritaikymas ir apmokymas (2)

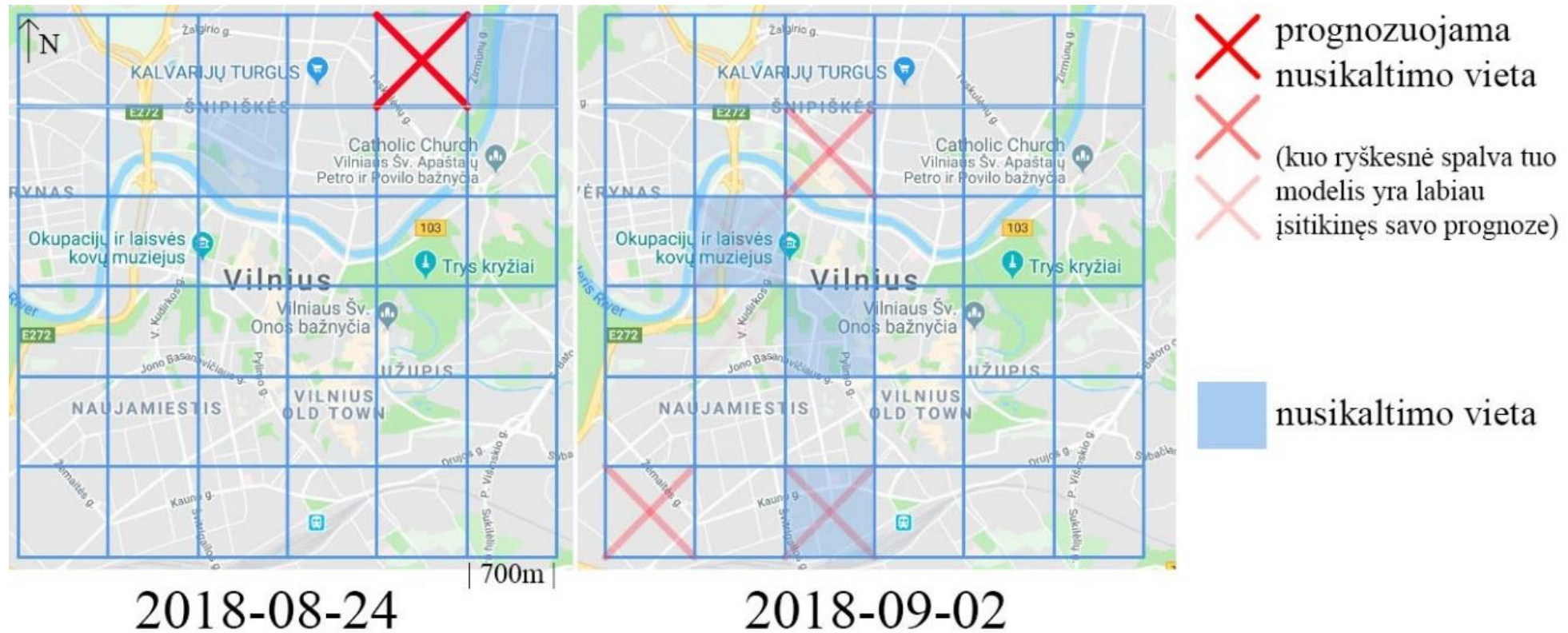


4 pav. Modelio prognozavimo rezultatai su skirtingais apmokymo parametrais



5 pav. Modelio prognozavimo rezultatai be aplinkos duomenų ir su aplinkos duomenimis

Rezultatai (1)



6 pav. LSTM modelio prognozės rezultatai su testavimo duomenimis

Rezultatai (2)

1. Buvo surinkti nusikaltimų ir aplinkos duomenys.
2. Buvo pasirinkta LSTM modelio prognozės vertinimo strategija.
3. Buvo sukurtas modelis kuris prognozuoja nusikaltimų vietas kitai dienai.
4. Papildomų aplinkos duomenų panudojimas modelio apmokymui modelio prognozės rezultatų nepagerino

Tyrimo rezultatai buvo pristatyti „Lietuvos magistrantų informatikos ir IT tyrimai“ konferencijoje

Išvados

1. Papildomų aplinkos duomenų panudojimas modelio apmokymui modelio prognozės rezultatų nepagerino, nes aplinkos duomenys buvo statiniai.
2. Patyrusiems policininkams, tokie LSTM modelio prognozavimo rezultatai gali suteikti naujų įžvalgų, kurių policijos pareigūnai galbūt nežinojo ar neįvertino.

Tolimesni darbo plėtojimai:

- Daugiau papildomų duomenų.
- Panaudoti kitą grupavimo algoritmą.
- Tirti kitus nusikaltimų tipus.
- Apmokymų duomenų skaidymas.
- Panaudoti kitą LSTM modelio architektūrą.