

Rozpoznávanie gest pomocou strojového učenia

Matej Martinček
Školiteľ: Ing. Viktor
Kocur, PhD.



Ciele práce

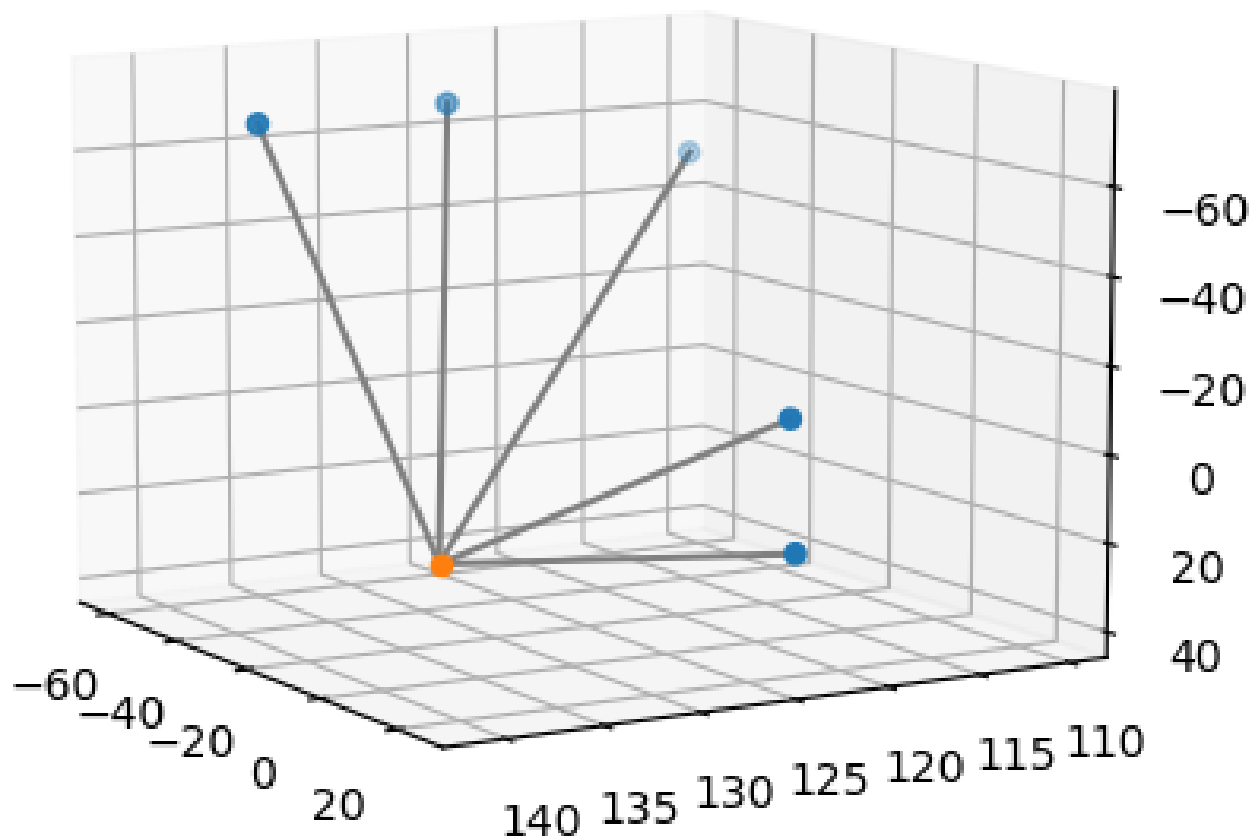
- Rozpoznávanie gest – metódy strojového učenia
- Porovnanie štandardne používaných klasifikátorov a nového modelu neurónovej siete
- Nadviazanie na pilotnú štúdiu:
 - modul pre rozpoznávanie gest
 - humanoidný robot NICO

Dataset



LEAP MOTION CONTROLLER (LMC)

- Zaznamenávanie pohybu ruky
- Infračervené kamery a LED svetlá
- Záznamy: časová sekvencia snímok (snímka=pozícia ruky (jej častí))

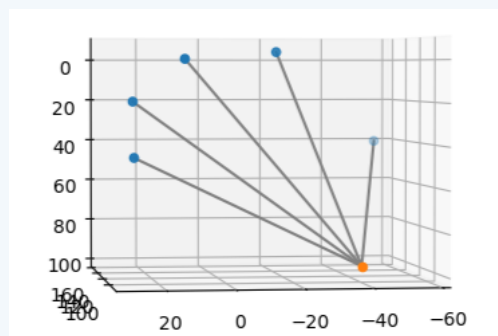
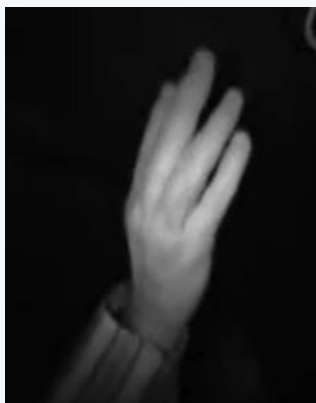


Naše dáta

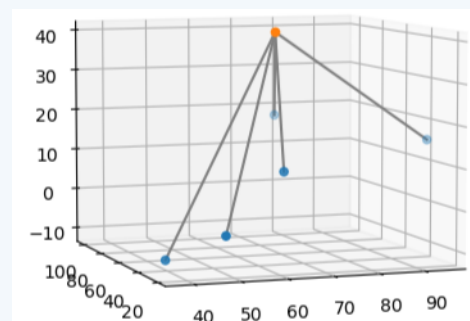
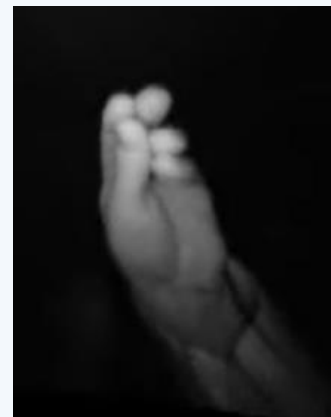
- Jedna snímka = pozícia dlane a koncov prstov = vektor dĺžky 18
- 6 typov záznamov (gest) / klasifikačných tried

Ukážka rozpoznávaných gest a vizualizácia vektora popisujúceho snímku daného typu gesta

1. pýtať si

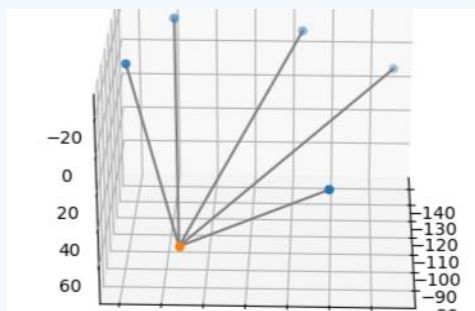


2. uchopiť

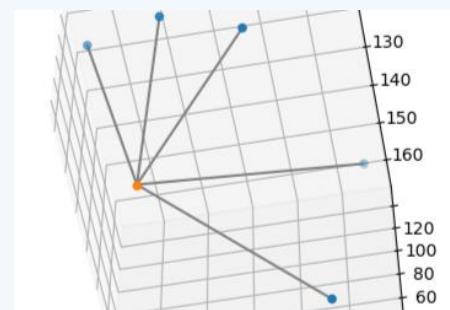
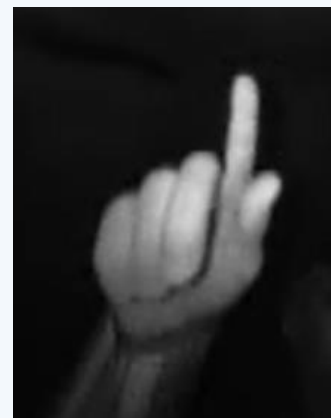


Ukážka rozpoznávaných gest a vizualizácia vektora popisujúceho snímku daného typu gesta

3. mávať

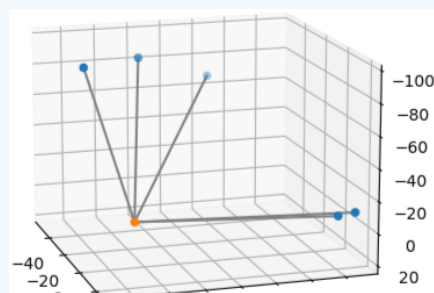
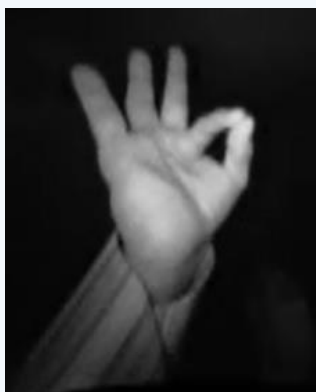


4. ukázať

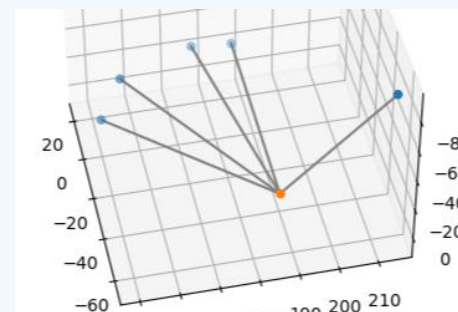


Ukážka rozpoznávaných gest a vizualizácia vektora popisujúceho snímku daného typu gesta

5. ok



6. nič



Ďalšie parametre datasetu

- 11 subjektov
- Jeden záznam gesta = niekoľko (5-10) opakovaní gesta (= 60 - 200 snímok)
- Počet záznamov: 192 => 32 pre každý typ gesta
- Počet snímok: 22 000

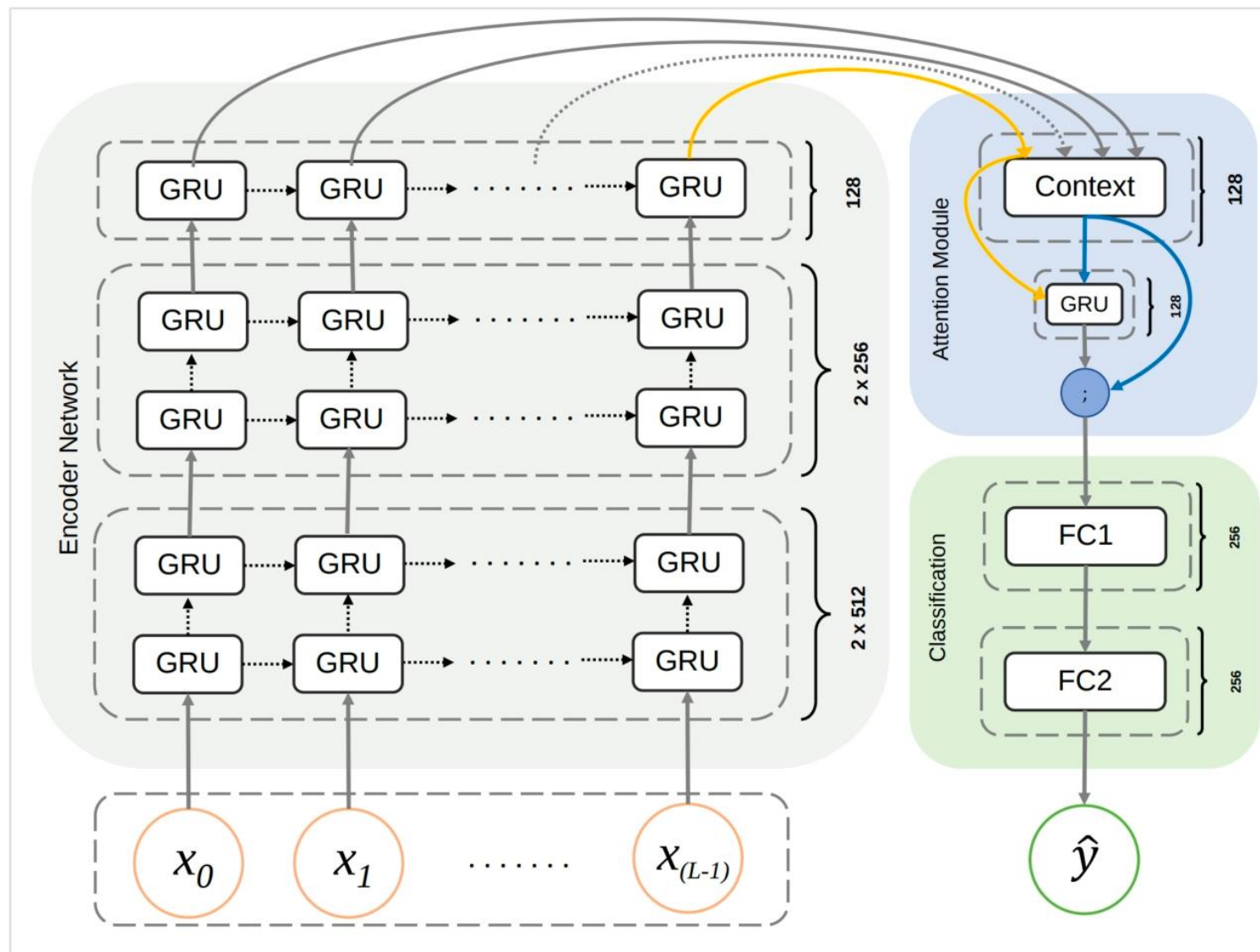
Experimenty a výsledky

Porovnanie klasifikačných metód a predspracovania dát

Klasifikačné metódy

1. Metóda podporných vektorov (SVM)
2. Náhodný les
3. Neurónová sieť DeepGRU

Neurónová sieť DeepGRU



Predspracovanie dát

1. Vlastné predspracovanie
2. Tvorba sekvencie snímok zodpovedajúcej jednému gestu
3. Škálovanie
4. PCA (analýza hlavných komponentov)

Vlastné predspracovanie

- Aplikované na jednotlivé snímky (reprezentované vektorom dĺžky 18)
 1. **Centrovanie:** Odpočítanie pozície stredu dlane od pozícií prstov
 2. **Normalizácia:** Predelenie centrovaných pozícií prstov vzdialenosťou prostredníka od stredu dlane
- Skrátene: „center-norm“ predspracovanie

Vlastné predspracovanie

Formálne:

\vec{d} - dlaň

\vec{p} - prostredník

Pôvodná snímka (ako zretáženie vektorov): $(\vec{d}, \vec{a}, \vec{b}, \vec{p}, \vec{c}, \vec{e})$

Pre každé $\vec{x} \in \{\vec{a}, \vec{b}, \vec{p}, \vec{c}, \vec{e}\} \subseteq \mathbb{R}^3$:

$$\vec{x}^* = \frac{\vec{x} - \vec{d}}{|\vec{p} - \vec{d}|}$$

Nová snímka: $(\vec{d}, \vec{a}^*, \vec{b}^*, \vec{p}^*, \vec{c}^*, \vec{e}^*)$

Tvorba sekvencie snímok zodpovedajúcej jednému gestu – SVM a náhodný les

- Sekvencie ako vektory
- Testované dĺžky sekvencií: 1,10,20,40

Tvorba nových dát pre sekvenciu danej dĺžky:

- pre každú súvislú podpostupnosť (danej dĺžky) snímok záznamu vytvor zretážením nový vektor
- Príklad:
 - Dĺžka sekvencie: 2
 - Záznam: $\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_5, \vec{s}_6$
 - Nové dáta (ako vektory): $(\vec{s}_1, \vec{s}_2), (\vec{s}_2, \vec{s}_3), (\vec{s}_3, \vec{s}_4), (\vec{s}_4, \vec{s}_5), (\vec{s}_5, \vec{s}_6)$

Tvorba sekvencie snímok zodpovedajúcej jednému gestu – DeepGRU

- Sekvencie ako postupnosti vektorov

Tvorba nových dát:

- Rozdelenie každého záznamu gesta na sekvencie päťinovej dĺžky
- Príklad:
 - Záznam: $\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_5, \vec{s}_6, \vec{s}_7, \vec{s}_8, \vec{s}_9, \vec{s}_{10}$
 - Nové dáta (ako postupnosti vektorov): $(\vec{s}_1, \vec{s}_2), (\vec{s}_3, \vec{s}_4), (\vec{s}_5, \vec{s}_6), (\vec{s}_7, \vec{s}_8), (\vec{s}_9, \vec{s}_{10})$

Škálovanie dát

- Štandardné škálovanie – odpočítanie priemeru a vydelenie štandardnou odchýlkou
- Min-max škálovanie – škálovanie príznakov do rozsahu [0,1]

Postup:

1. Výpočet hodnôt potrebných premenných škálovača (napr. priemer/minimum/...) pre každý príznak osobitne cez všetky vstupné vektory
2. Aplikácia škálovania na každý príznak

Analýza hlavných komponentov

- Redukcia dimenzionality dát
- Testované počty ponechaných komponentov zachovávajúce 50%, 80%, 90%, 95%, 98%, 99% až 100% variácie v dátach
- Pre každú dĺžku sekvencie: odlišné dĺžky vstupných vektorov => odlišné testované počty ponechaných hlavných komponentov

Postup:

1. Analýza hlavných komponentov cez všetky vstupné vektory
2. Projekcia vstupných dát na nájdené komponenty

Vykonané experimenty

1. Výber kernelu pre SVM
 2. Výber optimálnej dĺžky sekvencie / predspracovania
 3. Výber zvyšných hyperparametrov modelov
- Implementačné detaily:
 - 5-fázová krížová validácia
 - mriežkové prehľadávanie (angl. gridsearch)
 - metrika: presnosť (percento správnych predikcií)

Výber kernelu SVM

- Fixná dĺžka sekvencie snímok gesta: 20
- Fixné hyperparametre modelu SVM
- Optimálny kernel: **RBF**

	Bez predspracovania		Center-norm predspracovanie	
	MinMaxScaler	StandardScaler	MinMaxScaler	StandardScaler
	Počet komponentov = 42			
linear	0.717087	0.712366	0.737352	0.73027
poly	0.719512	0.72338	0.745276	0.740422
rbf	0.78829	0.790323	0.784227	0.780293
	Počet komponentov = 360			
linear	0.727511	0.729805	0.732103	0.716563
poly	0.725936	0.728625	0.747965	0.749996
rbf	0.80127	0.80468	0.789995	0.785602
	Bez PCA			
linear	0.727577	0.729805	0.732037	0.716563
poly	0.772885	0.728625	0.770589	0.749996
rbf	0.797992	0.80468	0.795765	0.785602

Výber predspracovania a hyperparametrov – SVM a náhodný les

- Fixné hyperparametre SVM/náhodného lesa

Postup:

- Pre každú testovanú dĺžku sekvencie:
 1. Výber najlepšej kombinácie predspracovania testovaním všetkých možností (okrem PCA bez škálovania)
 2. Optimalizácia hyperparametrov klasifikátora pre najlepšiu kombináciu

Výber predspracovania a hyperparametrov – DeepGRU

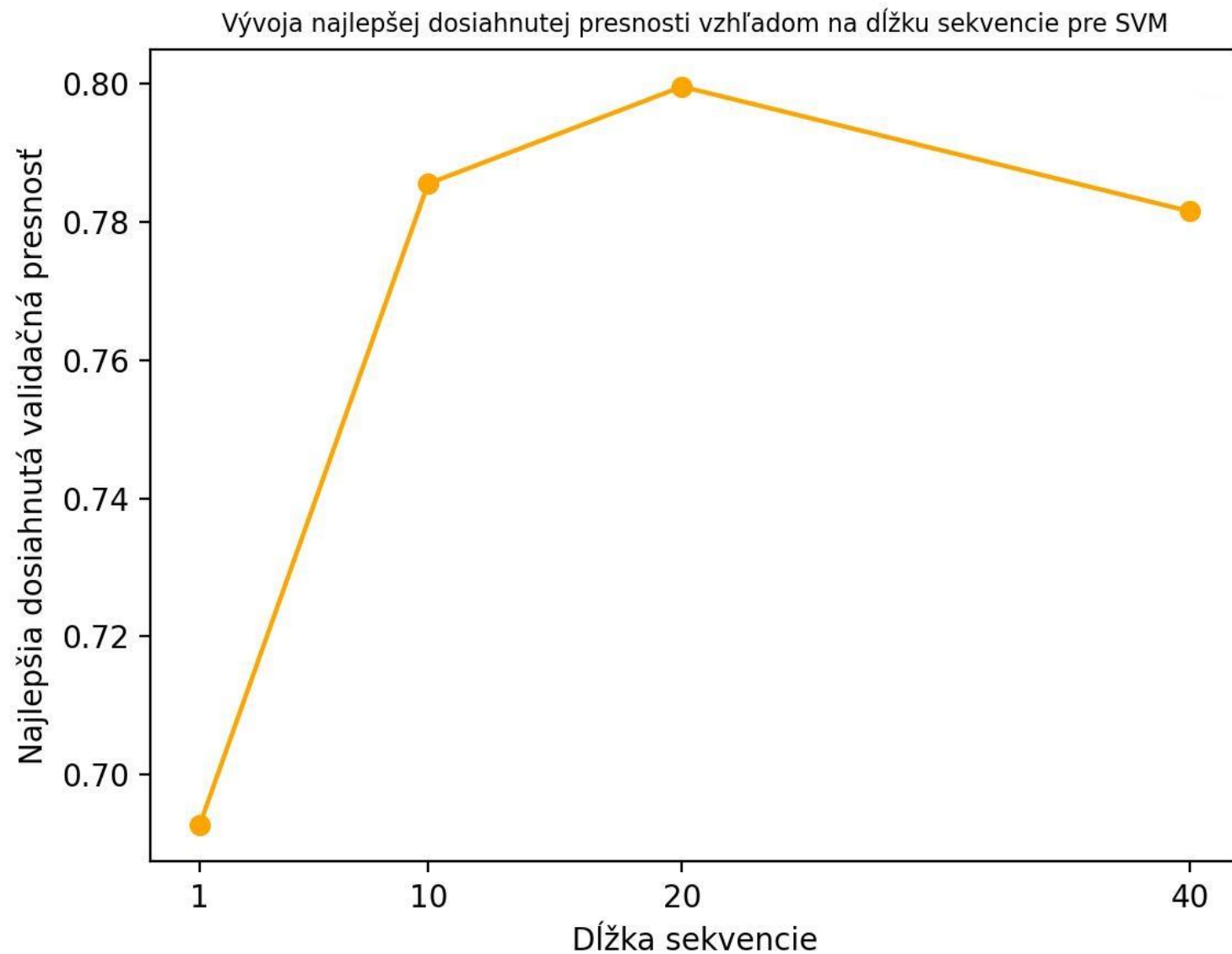
- Fixné hyperparametre (počet epoch, batchsize, ...)

Postup:

1. Výber najlepšej kombinácie predspracovania testovaním všetkých možností (okrem PCA bez škálovania)
2. Pre najlepšiu kombináciu predspracovania optimalizácia niektorých hyperparametrov

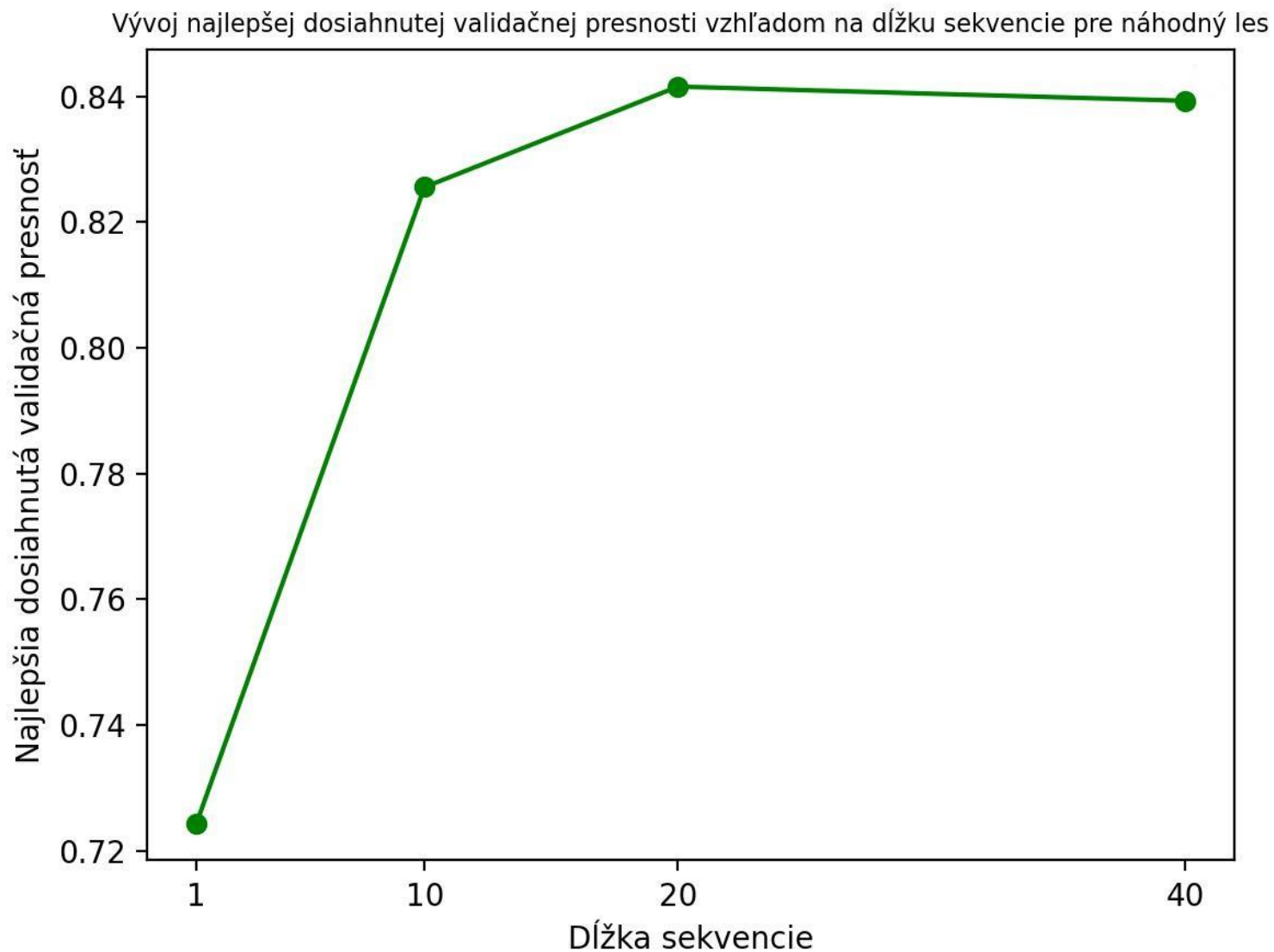
Optimálna dĺžka sekvencie – SVM

- Optimálna dĺžka sekvencie: 20



Optimálna dĺžka sekvencie – náhodný les

- Optimálna dĺžka sekvencie: 20



Optimálne predspracovanie - SVM

- Štandardné škálovanie bez vlastného predspracovania
- PCA: bez / ponechanie všetkých komponentov po PCA

	Bez predspracovania			Center-norm predspracovanie		
	min-max	štandardné	bez	min-max	štandardné	bez
bez PCA	0.797992	0.80468	0.791107	0.795765	0.785602	0.769738
PCA 5	0.383353	0.449709		0.623131	0.641031	
PCA 12	0.678794	0.69276		0.707776	0.729085	
PCA 23	0.774522	0.768097		0.772034	0.773411	
PCA 42	0.78829	0.789602		0.784752	0.780227	
PCA 79	0.79537	0.797337		0.785078	0.780226	
PCA 116	0.798385	0.801205		0.789274	0.784356	
PCA 360	0.80127	0.80468		0.789995	0.785602	

Optimálne predspracovanie – náhodný les

- Štandardné škálovanie s vlastným predspracovaním
- PCA: ponechaných 116 komponentov (99% variancie)

	Bez predspracovania			Center-norm predspracovanie		
	min-max	štandardné	bez	min-max	štandardné	bez
bez PCA	0.783436	0.781337	0.783042	0.785534	0.785075	0.78383
PCA 5	0.378564	0.445249		0.640638	0.667717	
PCA 12	0.635911	0.648371		0.681091	0.72404	
PCA 23	0.752624	0.757345		0.745739	0.77708	
PCA 42	0.779701	0.789996		0.760686	0.786125	
PCA 79	0.809993	0.822189		0.794585	0.805797	
PCA 116	0.819828	0.826123		0.804355	0.829204	
PCA 360	0.821205	0.814584		0.805209	0.820287	

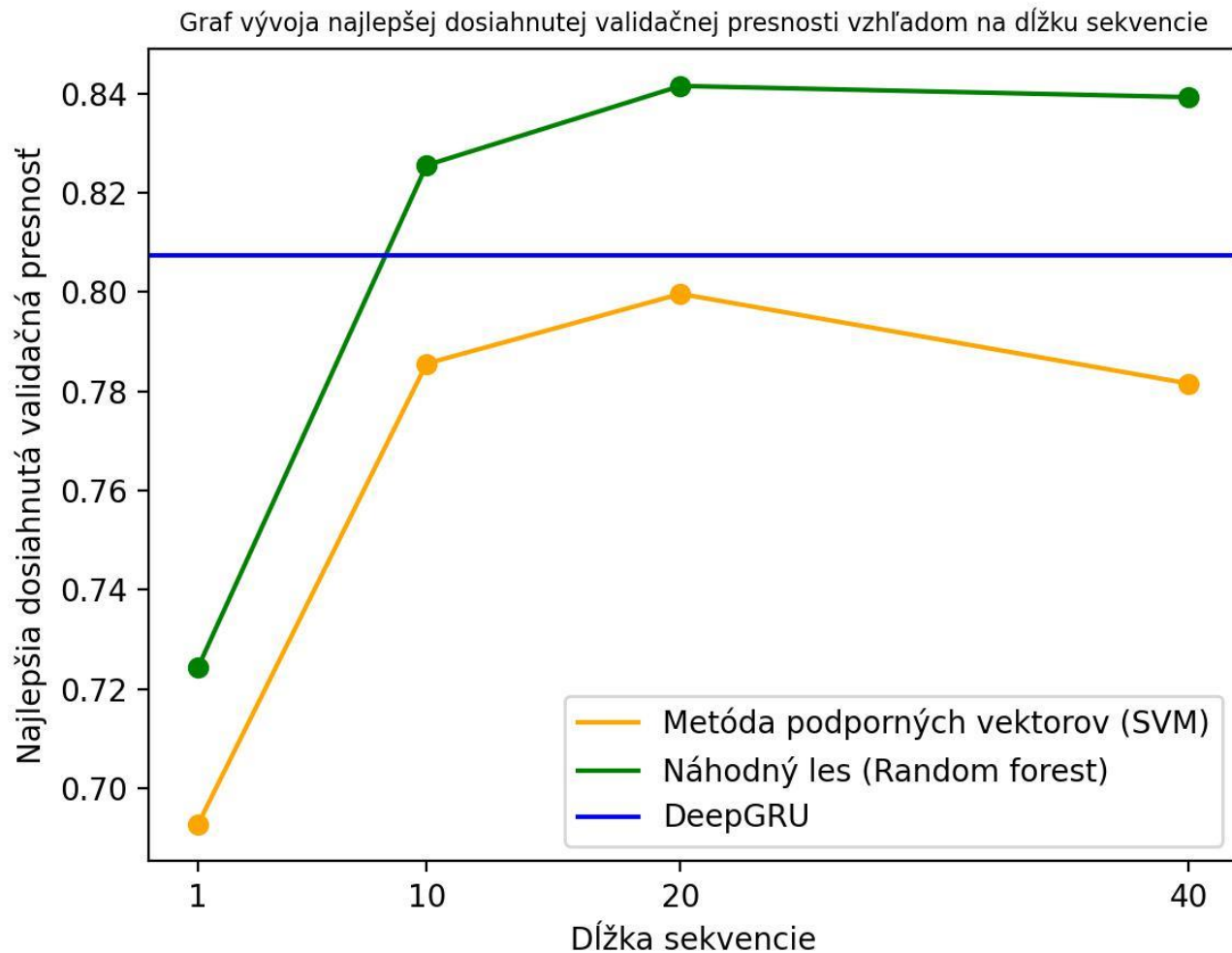
Optimálne predspracovanie – DeepGRU

- Štandardné škálovanie bez vlastného predspracovania
- PCA: ponechaných 12 komponentov (99% variancie)

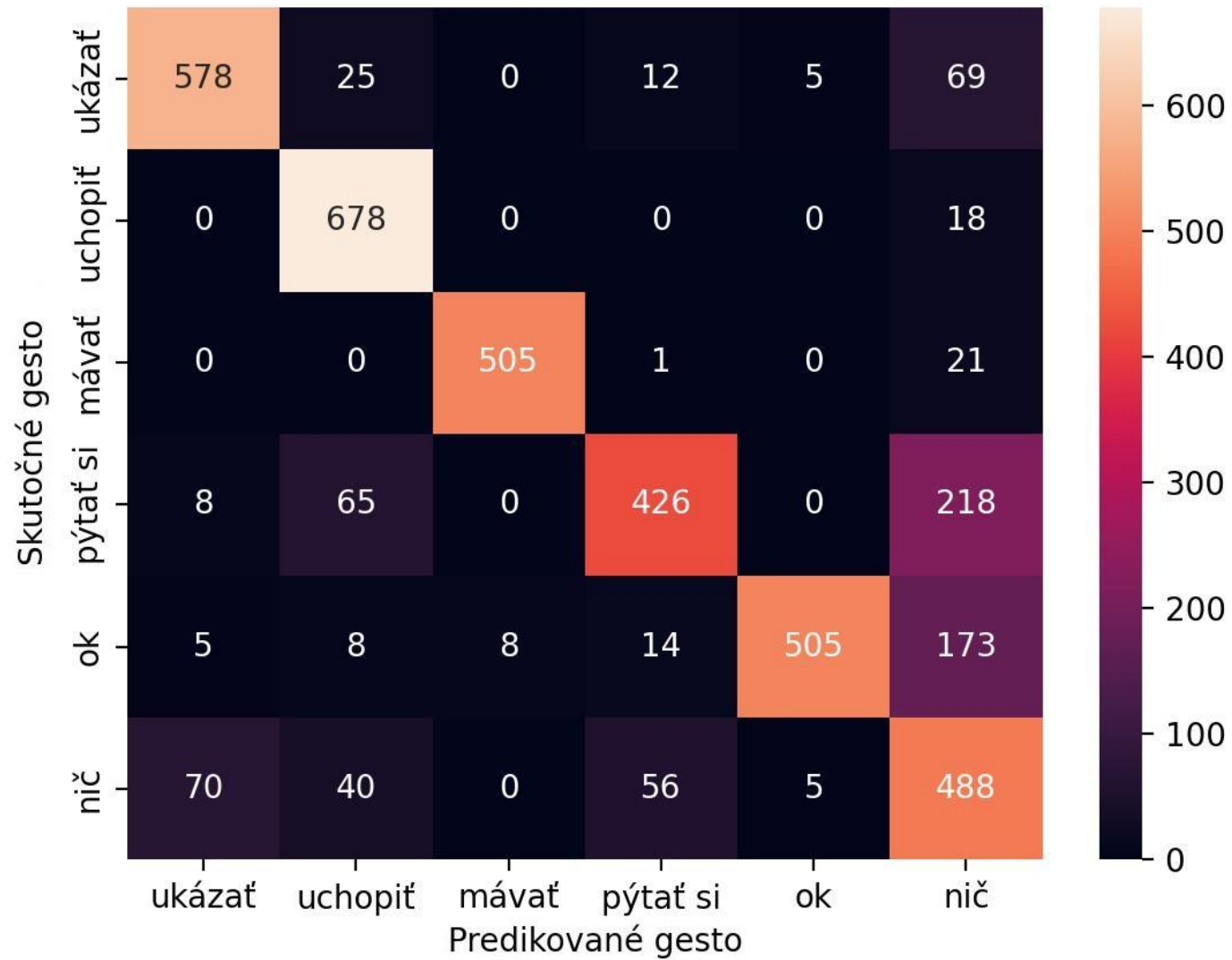
	Bez predspracovania			Center-norm predspracovanie		
	min-max	štandardné	bez	min-max	štandardné	bez
bez PCA	0.789881	0.772462	0.662891	0.700227	0.77703	0.758773
PCA 2	0.4858	0.502278		0.486859	0.52476	
PCA 4	0.478444	0.566398		0.614569	0.684219	
PCA 5	0.635068	0.714237		0.61459	0.706403	
PCA 28	0.687405	0.738595		0.698568	0.741147	
PCA 11	0.739863	0.806417		0.757348	0.793294	
PCA 12	0.74356	0.807468		0.759854	0.780857	
PCA 18	0.770091	0.801731		0.748754	0.790435	

Porovnanie klasifikátorov

- Najlepšia validačná presnosť (náhodný les): 84%
- Presnosť na testovacích dátach: 80%



Matica zámien



Záver

Celkový prínos

- Kód experimentov, ukladania/načítavania výsledkov a ich vizualizácie
- Návrh a implementácia vlastného predspracovania a tvorby sekvencií
- Celkový počet experimentov: **444**
- Počet natrénovaných modelov: **2 220**

Celkový prínos

- Dosiahnuté závery na datasete:
 - Najvhodnejší kernel pre SVM: **RBF kernel**
 - Približne optimálna dĺžka sekvencie snímok: **20**
 - Vhodnosť použitia PCA + škálovania dát (+ 3-5%)
 - Najvhodnejší klasifikátor: **náhodný les**
- Pokrok v presnosti na datasete o **10%** (80%)

Rozšírenie do budúcnosti

- Aplikácia dosiahnutých výsledkov pre rozpoznávanie gest v reálnom čase a tiež ako modul pre humanoidného robota NICO
- Návrh vlastnej neurónovej siete pre účely rozpoznávania gest (z tohto datasetu)



Ďakujem za
pozornosť!

Otázky z oponentského posudku

1. Prečo ste vybrali práve modely SVM, náhodný les a DeepGRU? Skúmali ste použitie aj iných modelov?
2. Zvažovali ste úpravu parametrov DeepGRU, ako sú počty neurónov vo vrstvách, prípadne zaradenie skorého zastavovania (early stopping)?
3. Zaoberali ste sa možnou ablačnou štúdiou? Myslíte, že by priniesla zaujímavé výsledky? (Pozn. ablačnou štúdiou sa označuje sada experimentov, pri ktorej prebieha odstraňovanie jednotlivých častí modelu, predspracovania, prípadne vstupných dát. Následne sa skúma vplyv týchto zmien na výkon modelu.)