

Používateľská príručka

pre aplikáciu k diplomovej práci

„Predikcia teploty a uhlíka v procese výroby ocele na báze strojového učenia“

Obsah

Systémové požiadavky	3
Inštalácia Python knižníc.....	3
Zoznam použitých knižníc	3
Prehľad aplikácie.....	4
Navigačné menu	4
Grafická časť.....	4
Konzola.....	4
Predspracovanie dát	5
Predikcia na statických dátach - statické dáta	5
Predikcia na statických dátach - cieľové dáta	5
Predikcia na zmiešaných dátach - statické dáta	5
Predikcia na zmiešaných dátach - dynamické dáta.....	6
Práca s aplikáciou.....	7
Načítanie dát.....	7
Nastavenia	8
Všeobecné nastavenia	8
Nastavenia pozorovaní.....	8
Nastavenia výstupov	9
Nastavenia modelu	9
Predikcia.....	10
Nastavenia grafickej časti.....	11
Export dát.....	11
Info	11

Systémové požiadavky

Aplikácia bola vyvíjaná pre operačné systémy Windows 10 a vyššie v prostredí Python verzia 3.8 a vyššie. Kompatibilita s inými operačnými systémami a staršími verziami nie je zaručená. Celková veľkosť aplikácie je približne 2 MB (bez virtuálneho prostredia).

Inštalácia Python knižníc

Pred prvým spustením aplikácie je potrebné nainštalovať používané knižnice nasledovným spôsobom:

1. V hlavnom priečinku aplikácie otvorte terminál, ktorý vie pristupovať k prostrediu Python (príkazový riadok, PowerShell a pod.). Ideálne by to malo byť virtuálne prostredie Python.
2. Spustíte príkaz `pip install -r requirements.txt`
3. Pomocou príkazu `python Predikcia.py` spustíte aplikáciu.

Zoznam použitých knižníc

Nasledujúci zoznam obsahuje všetky knižnice potrebné pre fungovanie aplikácie aj s uvedenou verziou, ktorá bola použitá pri vývoji aplikácie.

- `certifi==2025.1.31`
- `charset-normalizer==3.4.1`
- `contourpy==1.3.1`
- `cycler==0.12.1`
- `defusedxml==0.7.1`
- `et_xmlfile==2.0.0`
- `fonttools==4.55.8`
- `idna==3.10`
- `joblib==1.4.2`
- `kiwisolver==1.4.8`
- `matplotlib==3.10.1`
- `numpy==2.2.4`
- `packaging==24.2`
- `pandas==2.2.3`
- `pillow==11.1.0`
- `pyparsing==3.2.1`
- `PySide6==6.9.0`
- `PySide6_Addons==6.9.0`
- `PySide6_Essentials==6.9.0`
- `python-calamine==0.3.2`
- `python-dateutil==2.9.0.post0`
- `pytz==2025.1`
- `requests==2.32.3`
- `scikit-learn==1.6.1`
- `scipy==1.15.1`
- `seaborn==0.13.2`
- `shiboken6==6.9.0`
- `six==1.17.0`
- `threadpoolctl==3.5.0`
- `tzdata==2025.1`
- `urllib3==2.3.0`

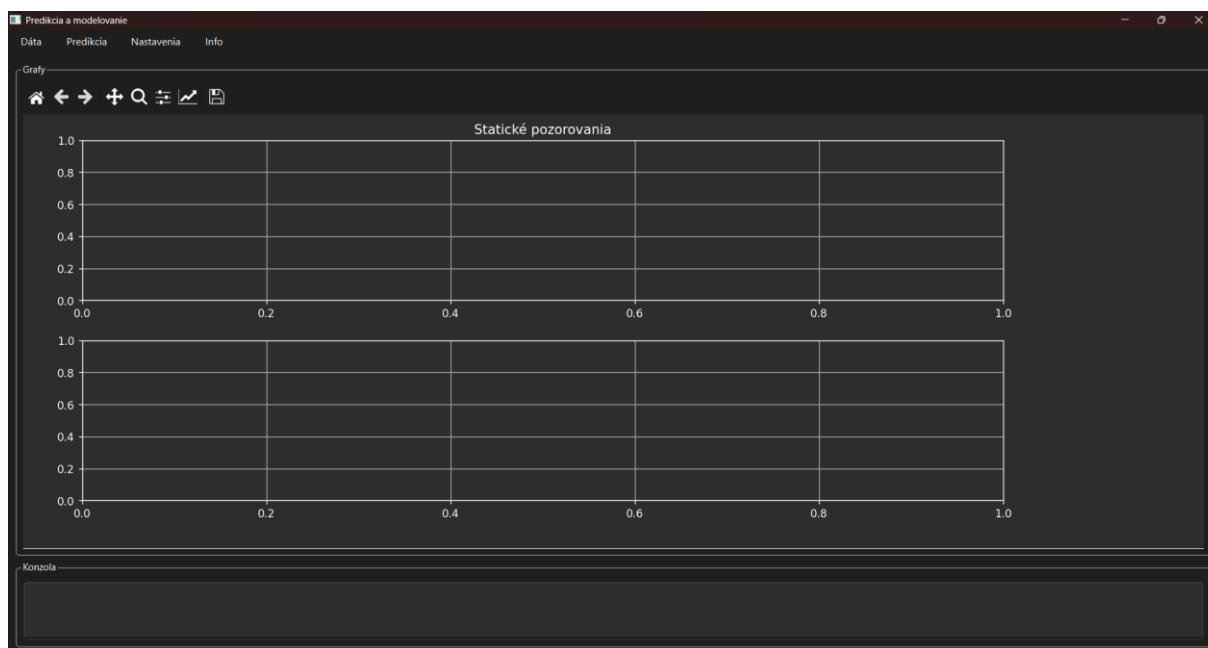
Odkaz na GitHub:

<https://github.com/BlueScreenOfREKT/predikcia-teploty-uhlika>

Prehľad aplikácie

Okno aplikácie (Obr. 1) je rozdelené na 3 hlavné časti a sú nimi:

- navigačné menu
- grafická časť
- konzola



Obr. 1

Navigačné menu

- Podmenu **Dáta** obsahuje položky, ktoré slúžia na načítanie a export dát.
- Tlačidlo **Predikcia** spúšťa predikciu na základe načítaných dát a zvolených nastavení.
- Podmenu **Nastavenia** obsahuje odkazy na nastavenia pozorovaní, cieľov predikcie či samotného modelu a iných.
- Podmenu **Info** obsahuje odkazy na základné informácie o aplikácii či používateľskú príručku.

Grafická časť

Je zložená z panela nástrojov v hornej časti a z figúry, ktorá je ďalej rozdelená na dva samostatné grafy. V hornom grafe sa vždy zobrazujú načítané statické pozorovania a v spodnom grafe sa v závislosti od typu žiadanej predikcie (načítaných dát) zobrazí buď žiadaný cieľ predikcie pri predikcii na statických dátach alebo dynamické pozorovania pri predikcii na zmiešaných dátach (statické aj dynamické).

Konzola

Slúži na výpis správ aplikácie. Jedná sa najmä o dôležité správy o chovaní aplikácie, výpis zmeny nastavení, oznamovanie chýb či výsledkov predikcie. Používateľ sa môže posúvať medzi jednotlivými správami a má tak prístup k histórii správ.

Predspracovanie dát

Na to aby bola možná predikcia musia byť dáta predspracované do vhodného formátu tak aby ich vedela aplikácia rozpoznať. Dáta sú vždy rozdelené na dva celky a to statické pozorovania a dynamické pozorovania alebo cieľ predikcie. Každý databázový súbor musí obsahovať tzv. hlavičku čiže prvý riadok, ktorý obsahuje názvy stĺpcov (názov osi x, názvy veličín a pod.). Samotné dáta potom musia obsahovať číselné hodnoty z oboru reálnych čísel. Aplikácia nespracuje dáta obsahujúce iné hodnoty (napr. NaN). V prípade potreby tiež viete nastaviť v aplikácii typ kódovania znakov.

Predikcia na statických dátach - statické dáta

Aplikácia očakáva jeden stĺpec, ktorý bude obsahovať por. čísla tavieb (predvolene *Tavba č.*) a súbor pozorovaní (ostatných stĺpcov). Očakávaný názov stĺpca por. čísel tavieb je však možné zmeniť priamo v nastaveniach aplikácie. Namerané hodnoty rovnakej veličiny ale v inom časovom úseku sú uložené v samostatných stĺpcoch v riadku príslušnej tavby (Obr. 2).

	A	B	C	D
1	Tavba č.	Sur. Fe - Al	02 - Skut. r	02 - Trv. t
2	1	684	8387.663	1088
3	2	684	8210.206	1011

Obr. 2

Predikcia na statických dátach - cieľové dáta

Aplikácia očakáva jeden stĺpec, ktorý bude obsahovať por. čísla tavieb (predvolene *Tavba č.*) a stĺpce s nameranými hodnotami teploty (predvolene *Teplota (°C)*) a koncentrácie uhlíka (predvolene *Uhlík (%)*) v konkrétnom časovom úseku (napr. koniec alebo začiatok tavby). Očakávaný názov stĺpca por. čísel tavieb je však možné zmeniť priamo v nastaveniach aplikácie (Obr. 3).

	A	B	C	D
1	Tavba č.	Teplota (°C)	Uhlík (%)	
2	1	1672	0.027	
3	2	1644	0.026	

Obr. 3

Predikcia na zmiešaných dátach - statické dáta

Aplikácia očakáva dáta získané v rôznych časových úsekoch tavby najmä začiatok a koniec pričom namerané hodnoty jednotlivých veličín sú v jednom stĺpci za sebou tak ako boli namerané. Merania musia byť pre všetky veličiny vykonané v rovnakých časových úsekoch (Obr. 4).

	A	B	C
1	Teplota (°C)	Uhlík (%)	CO v spal
2	1243.858	4.32	
3	1230.264	4.247	1.12847

Obr. 4

Predikcia na zmiešaných dátach - dynamické dáta

Aplikácia očakáva jeden stĺpec, ktorý bude obsahovať časovú postupnosť (predvolene *Čas (s)*) a súbor pozorovaní, ktorých názvy stĺpcov sa zhodujú s názvami stĺpcov v databáze statických dát. Predikčný model vie pracovať len s pozorovaniami, ktoré sú spoločné pre obe databázy (Obr. 5).

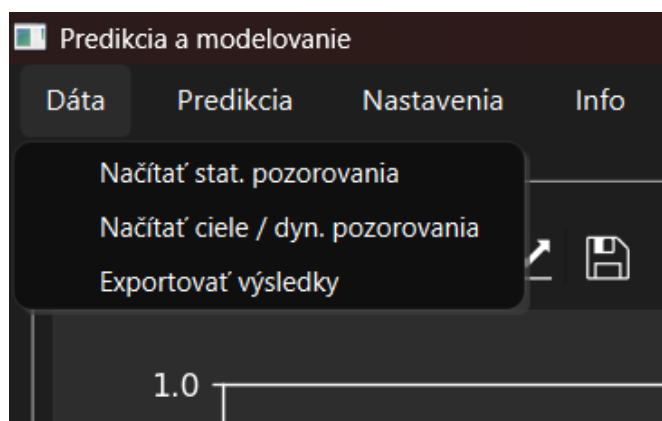
	A	B	C
1	Čas (s)	CO v spal.	CO2 v spa
2	0	0	1.93865
3	1	0	1.93865

Obr. 5

Práca s aplikáciou

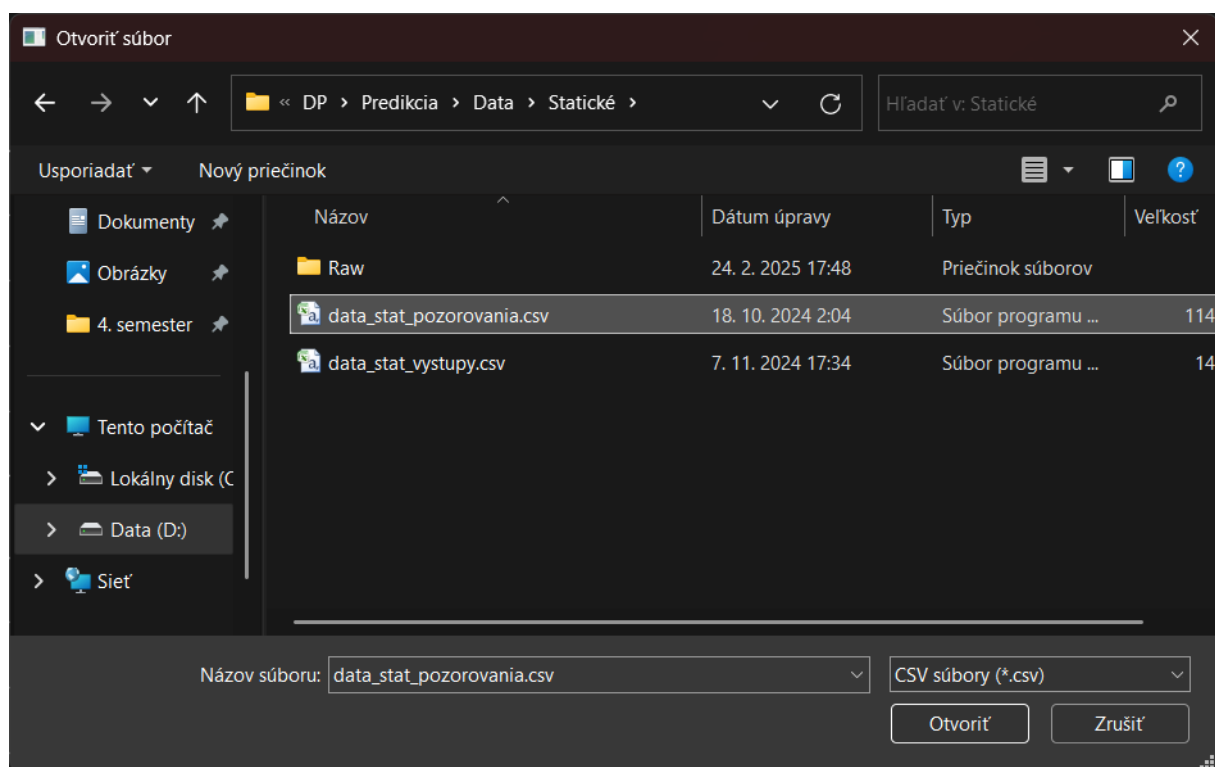
Načítanie dát

1. V navigačnom menu (Obr. 6) zvolíte **Dáta** a následne vyberiete typ dát, ktorý chcete aktuálne načítať.



Obr. 6

2. Otvorí sa vám okno v ktorom môžete vybrať databázový súbor. V prípade, že súbor nevidíte skontrolujte máte zvolený správny typ súboru prípadne či je typ vášho databázového súboru podporovaný (Obr. 7).

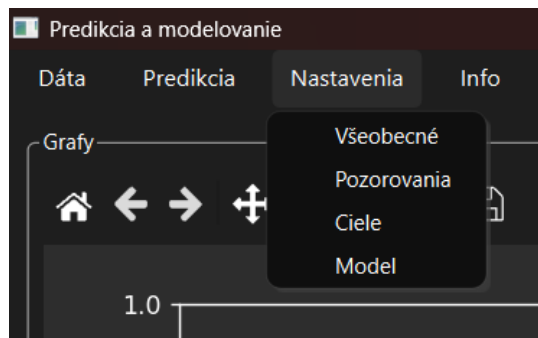


Obr. 7

3. Ak sú dáta v požadovanom formáte, vykreslia sa v zodpovedajúcom grafe. Inak sa vypíše chybové hlásenie.

Nastavenia

Cez podmenu **Nastavenia** (Obr. 8) prístupujete k rôznym nastaveniam aplikácie.

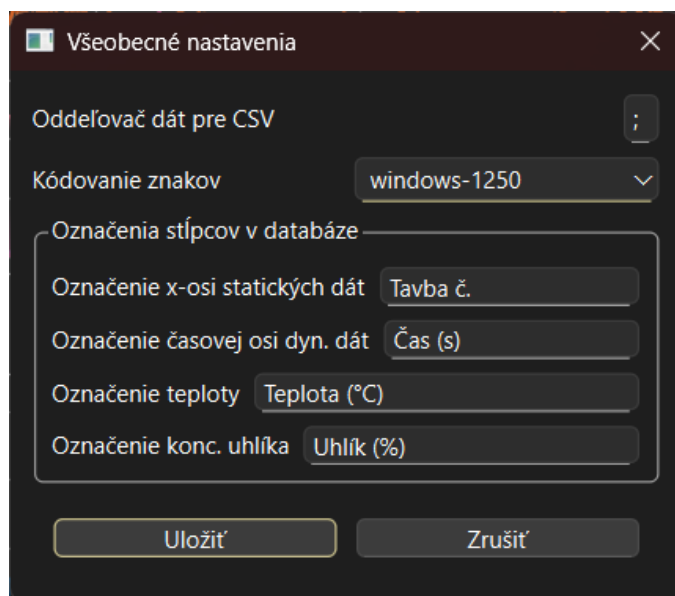


Obr. 8

Všeobecné nastavenia

V menu **Všeobecné nastavenia** (Nastavenia > Všeobecné) sú dostupné nasledujúce možnosti (Obr. 9):

- Oddeľovač dát pre CSV – nastavte znak, ktorý používate vo vašom databázovom súbore na oddelenie jednotlivých hodnôt
- Kódovanie znakov – tu môžete nastaviť kódovanie vášho databázového súboru (vyberte z predvolených možností alebo napíšte názov iného kódovania, ktoré podporuje funkcia *pandas.read_csv*).
- Označenia stĺpcov v databáze – pre správnu funkciu aplikácie sú potrebné správne označenia kritických stĺpcov v databáze. Na Obr. 9 sú znázornené preddefinované názvy, ktoré môžete zmeniť tak aby sa zhodovali s vašim databázovým súborom.



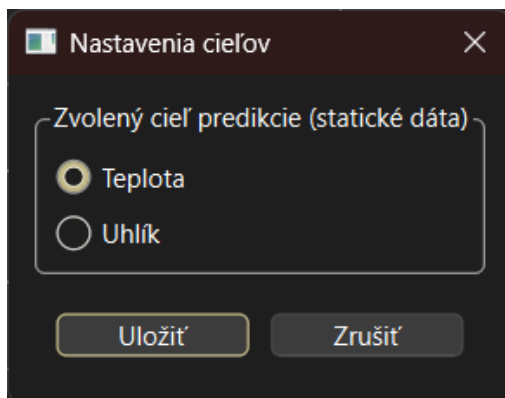
Obr. 9

Nastavenia pozorovaní

V menu **Nastavenie pozorovaní** (Nastavenia > Pozorovania) môžete zvoliť, ktoré pozorovania chcete vykresliť v grafe a zahrnúť v predikcii. Po uložení výberu sa graf automaticky aktualizuje. Pri predikcii na zmiešaných dátach budú však vždy zvolené len tie pozorovania, ktoré sú spoločné pre obe databázy (statickú aj dynamickú).

Nastavenia výstupov

V menu **Nastavenia cieľov** (Nastavenia > Ciele) môžete zvoliť aktuálny cieľ predikcie pri predikcii zo statických dát (Obr. 10).



Obr. 10

Nastavenia modelu

V menu **Nastavenia modelu** (Nastavenia > Model) je viacero nastavení súvisiacich s predikčnými modelmi.

Nastavením pomeru trénovacích dát zmeníte aká časť dát (v percentách) sa použije na trénovanie modelu pri predikcii zo statických dát. Zvyšná časť dát sa použije na overenie výkonnosti predikcie natrénovaného modelu.

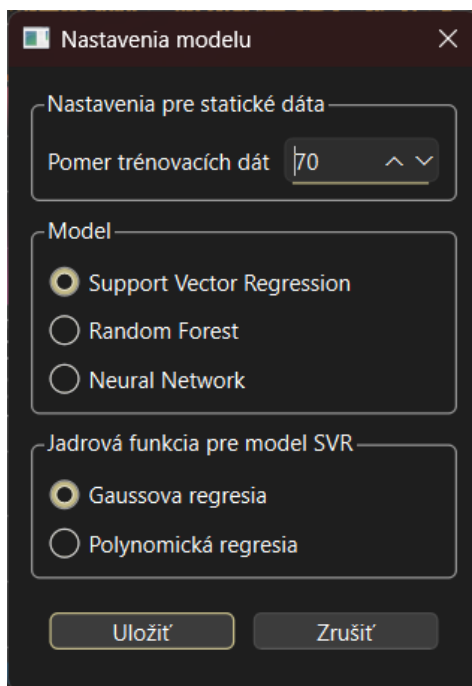
Pre akúkoľvek predikciu viete zvoliť použitý jeden z dostupných modelov predikcie. Dostupné sú 3 modely:

- Support Vector Regression (SVR)
- Random Forest (RF)
- Neural Network (NN)

Ak sú k zvolenému modelu dostupné aj ďalšie nastavenia, tieto sa zobrazia pod voľbou modelu.

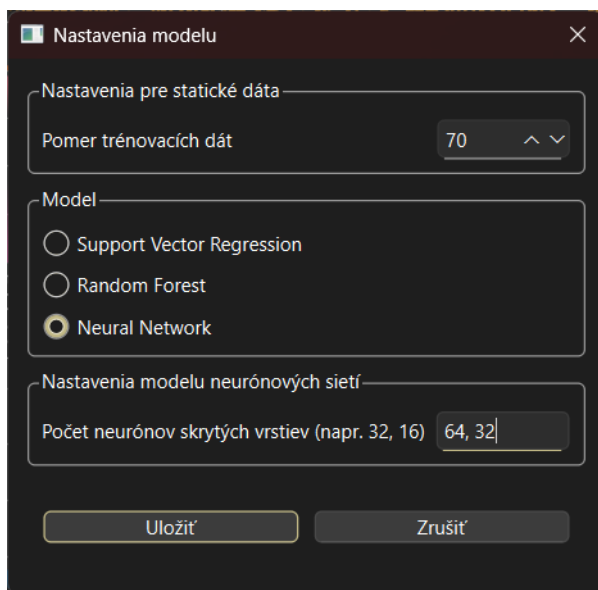
Pre **SVR** model (Obr. 11) je možné vybrať z dvoch jadrových funkcií:

- Gaussová - SVR(kernel='rbf')
- Polynomická - SVR(kernel='poly')



Obr. 11

Pre model **neurónových sietí** (Obr. 12) je možné zvoliť počet neurónov v jednotlivých skrytých vrstvách resp. počet skrytých vrstiev. Napríklad ak zadáte len číslo 50, predikčný model NN bude mať jednu skrytú vrstvu o veľkosti 50 neurónov. Ak zadáte 32, 16, 8 model bude obsahovať 3 skryté vrstvy – prvá vrstva bude mať 32 neurónov, druhá 16, tretia 8. **Upozorňujeme však, že s každým ďalším neurónom a vrstvou sa môže výrazne predĺžiť výpočtový čas.**



Obr. 12

Predikcia

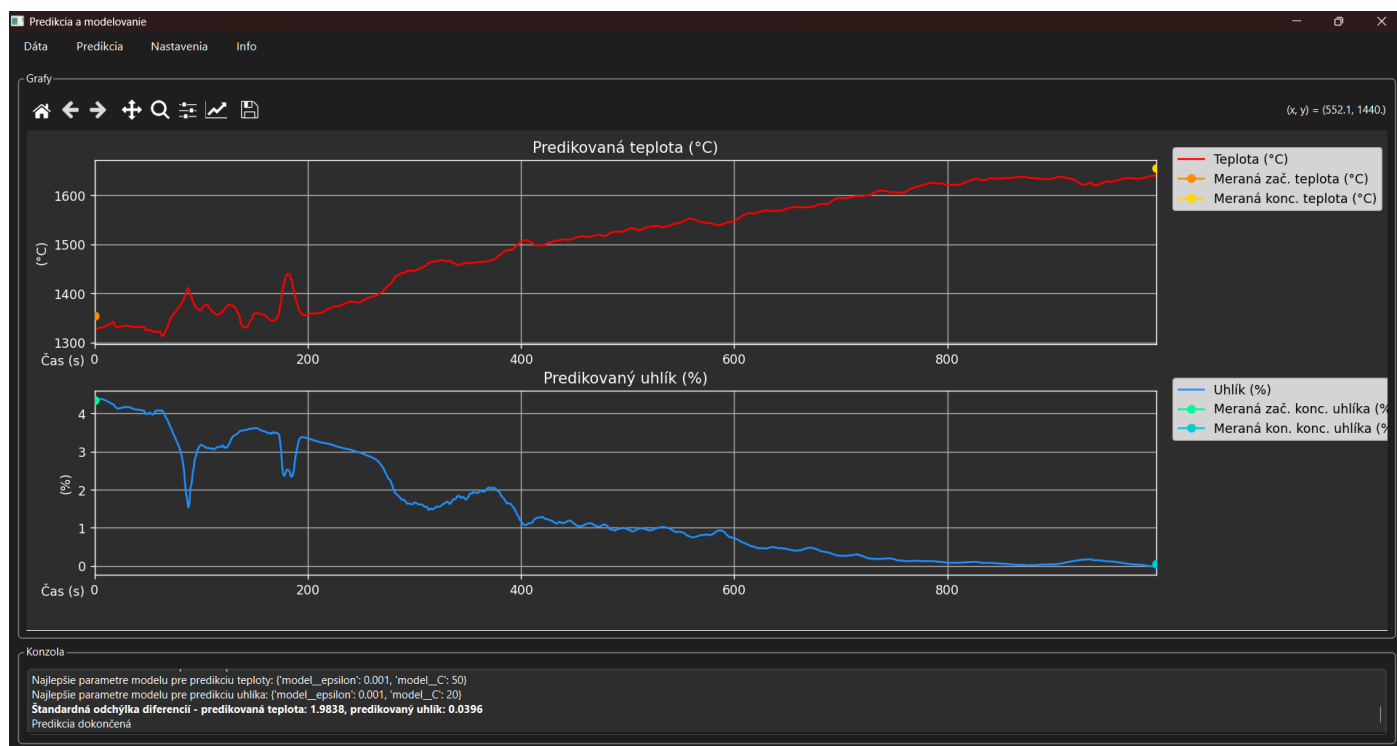
Tlačidlo **Predikcia** spustí predikciu s vybraným modelom a parametrami ak sú načítané a rozpoznané potrebné dáta. Ak nie sú dáta načítané, aplikácia o tom informuje používateľa v konzole.

Pri predikcii na zmiešaných dátach sa ešte pred samotným spustením výpočtového procesu zobrazí dialógové okno, kde môže používateľ zadať merané hodnoty teploty a konc. uhlíka v tavenine na začiatku aj na konci konkrétnej tavby. Ak ich používateľ nepozná, môže pokračovať bez ich zadania. Slúžia len na zobrazenie v grafe – vizuálne porovnanie.

Predikcia môže trvať v závislosti od množstva dát, zvoleného modelu, parametrov modelu a výpočtového výkonu zariadenia niekoľko sekúnd ale aj desiatky minút. Buďte opatrní a trpezliví.

V prípade predikcie pomocou neurónovej siete je automaticky zvolený počet neurónov skrytých vrstiev zadaných používateľom, ak ich zadal. Ak ich používateľ nezmenil v nastaveniach, použijú sa hodnoty s preddefinovanej množiny nastavení hyperparametrov.

Po skončení predikcie sa výsledky predikcie zobrazia v grafickej časti (Obr. 13). V konzole môžu byť vypísané štatistické údaje, ktoré možno použiť pre hodnotenie výkonnosti predikčného modelu a taktiež najlepšie nastavenia hyperparametrov modelu z preddefinovanej množiny nastavení (použitie *RandomizedSearchCV*).



Obr. 13

Nastavenia grafickej časti

Nad grafmi sa nachádzajú viaceré nástroje na prácu s grafmi zoradené tak ako za sebou nasledujú príslušné ikony:

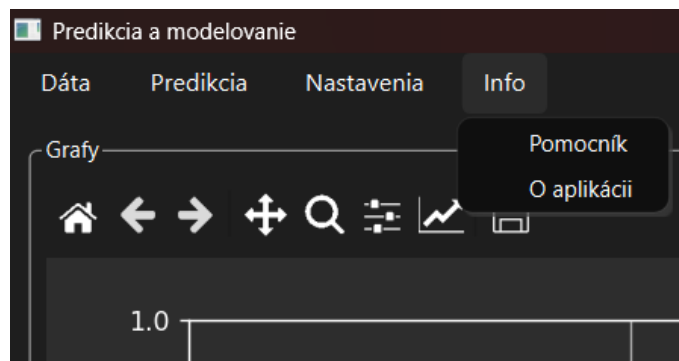
- obnovenie pôvodného zobrazenia
- prepínanie medzi poslednými zobrazeniami - dozadu
- prepínanie medzi poslednými zobrazeniami - dopredu
- posúvanie sa po grafe
- približovanie
- nastavenia jednotlivých grafov
- nastavenie jednotlivých osí
- export obrázkov grafickej časti

Export dát

Po zvolení možnosti **Exportovať výsledky** (Obr. 6) sa zobrazí okno, ktoré vám umožní zvoliť miesto uloženia výsledkov (predikovaných hodnôt) a názov súboru. Export výsledkov je vo formáte CSV.

Info

V podmenu **Info** (Obr. 14) sa nachádzajú položky **Pomocník** a **O aplikácii**. Prvá z uvedených zobrazí tento dokument a druhá krátke zhrnutie účelu aplikácie príp. iných informácií.



Obr. 14