



AI dětem

Kurikulum umělé inteligence pro ZŠ a SŠ

Karta

Velká data

01

balíček Data



kurikulum.aidetem.cz/informatika2

Vypracoval: Radek Špáta

Metodická konzultantka: Eva Nečasová

Odborní garanti: Tomáš Mlynář, Ondra Michalák a Pavel Kordík

Jazyková korektura: zatím neproběhla

Poslední aktualizace: 09/2024

Verze: 03



Národní pedagogický institut
České republiky

Metodiky vytváříme ve spolupráci s Národním pedagogickým institutem.



[Formulář pro
přípomínky.](#)

Metodický materiál Kurikula umělé inteligence pro základní a střední školy
Informatika na 2. stupni ZŠ a SŠ – karty

Velká data

Toto by bylo dobré vědět o velkých datech pro začátek

Ustálená definice pojmu velká data (Big Data) neexistuje, avšak společné charakteristiky velkých dat zahrnují jejich ohromnou velikost, to jak rychle vznikají, jak jsou různorodá (velikostí a strukturou) a také, že mají často nižší kvalitu nebo nejsou tak přesná. Velká data mohou zahrnovat obrázky, videa, audia, texty nebo tzv. digitální stopy ve formě údajů o uživatelském chování. Jsou důsledkem zrychlení a vývoje internetu, přičemž roli hraje také výrazné zlevnění ukládání a zpracování dat. Typicky je jich tolik, že jejich zpracování vyžaduje nové přístupy, například právě metody strojového učení.

A tohle, pokud vás to téma zajímá 🍷

Větší objemy dat začaly vznikat už zhruba v 80. letech minulého století [[zdroj](#)], ale velkými daty je nenazýváme. Velkou změnu kolem roku 2002 přinesly až systémy, které umožňovaly zpracovávat velmi rychle velké objemy dat na běžně dostupných počítačích, což výrazně zlevnilo celý proces a tím podpořilo také nárůst velkých dat.

Velká data pocházejí z různých zdrojů, jako jsou IoT zařízení, záznamy uživatelského chování na internetu, dopravní systémy, finanční transakce, zdravotnická data nebo vědecké experimenty. Jejich aplikace jsou velmi široké – od personalizace služeb a prediktivní analýzy ve zdravotnictví či financích po optimalizaci logistiky, zlepšení veřejných služeb a řízení chytrých měst. Kromě technologických společností využívají velká data také státní organizace, nemocnice a vědecké instituce a další.

Strojové učení hraje v oblasti velkých dat zásadní roli. Lidé mohou díky této metodě vytvářet modely, které nacházejí ve velkých datech vzory a na základě toho předpovídat budoucí vývoj situací, personalizovat různé produkty či služby a to mnohdy v reálném čase. Velká data se také využívají například pro trénování velkých jazykových modelů (LLM).

Balíček Data

Velká data

Velká data v čase
Vizualizace dat
Predikce z dat



[Prezentace k lekci v PDF](#)



[Editovatelná prezentace v Canva](#)

[Obecný úvod do AI](#)

Chcete se dozvědět více o umělé inteligenci? Připravili jsme pro vás srozumitelnou online příručku [Obecný úvod do umělé inteligence pro dospělé](#).

[Hledáte podporu?](#)

Nevíte si rady? Připojte se do [FCB skupiny AI dětem](#) a zeptejte se komunity nebo správců.

Pozn.: Genderová rovnost je pro AI dětem klíčová, ale pro zestručnění využíváme v našich metodikách formulace v mužském rodě.

Aktivita do hodiny

Vzory v lidském chování

30–45
minut

Popis aktivity

Cílem úvodní aktivity balíčku Data je uvědomění žáků, že velká data obsahují informace o chování lidí v běžném životě a lze z nich na základě analýzy (např. pomocí vizualizace pomocí grafů) interpretovat naše minulé chování, případně vytvářet předpovědi chování budoucího.

Žáci v aktivitě nejdříve pomocí aplikace The New Normal prozkoumávají vizualizace velkých dat z vyhledávače Google a porovnávají změny nákupního chování uživatelů na internetu. Poté pomocí stránek What are we searching for hledají vzory v lidském chování v různých regionech. A následně hledají a interpretují změny v chování uživatelů internetu před, během a po období pandemie Covid-19 v nástroji Google Trends.

Jak aktivita souvisí s velkými daty

Dle článku na stránce [seo.ai](#) uživatelé zadají do vyhledávače Google cca 99 000 vyhledávacích dotazů každou vteřinu, což znamená zhruba 8,5 miliard vyhledávacích dotazů každý den. Tato velká data jsou následně upravena do struktury, na jejímž základě nástroj Google Trends tato data vizualizuje. Velká data ovšem nevznikají pouze vyhledáváním na internetu, ale třeba také při používání sociálních sítí, záznamu transakcí v online obchodech a bankách a při sběru dat z chytrých zařízení, jako jsou senzory a nositelná elektronika (wearables).

Informace o aktivitě

Prekoncepty/ročníky, délka lekce

8. a 9. ročníky ZŠ a SŠ

Klíčové slovo

Velká data, Big Data

Co se žáci učí?

Velká data mohou být nástrojem pro porozumění jevům.

Proč se to učí?

Rozvíjí kritické myšlení a schopnost analýzy a interpretace komplexních jevů.

Jak poznáme, že se to naučili?

Analyzují a interpretují vzory v nákupních chování uživatelů internetu.

Pomůcky

Pedagog: Projekční zařízení, prezentace k promítnutí
Žáci: digitální zařízení do skupin nebo dvojic

Výstupy RVP – Informatika

I-9-1-01 získá z dat informace, interpretuje data, odhaluje chyby v cizích interpretacích dat

Digitální kompetence

Informace a komunikace: získává, vyhledává, kriticky posuzuje, spravuje a sdílí data, informace a digitální obsah, k tomu volí postupy, způsoby a prostředky, které odpovídají konkrétní situaci a účelu

Pět velkých myšlenek

3-A-II Podstata učení (hledání vzorů v datech)

Bloomova taxonomie

Porozumění: Žáci porovnávají a interpretují grafické vyjádření na základě velkých dat.

Analýza: Žáci analyzují vizualizace velkých dat.

Evokace

5
minut

Sdílejte

Promítněte prezentaci na straně O2.

Vzpomenete si, kdy Světová zdravotnická organizace vyhlásila epidemii Covid-19 za celosvětovou pandemii?

Odpověď: Bylo to 11. března 2020

Vybavte si začátek období pandemie. Co se ve vašem životě změnilo?

Možné odpovědi: Školy, obchody, restaurace a kina se zavřely. Lidé museli většinou zůstat doma a mohli jít ven jen na nákupy nebo na procházky. Museli jsme začít nosit roušky. Hodně států zavřelo hranice a lidé nemohli snadno cestovat do jiných zemí.

Změnily se nějakým způsobem vaše návyky? Vstávali jste například později? Jak vaši rodiče nakupovali? Potřebovali jste najednou jiné zboží, než před pandemií?

Možné odpovědi: Lidé vstávali později, protože nebyla škola a také nemohli chodit do práce. Rodiče často nakupovali online, protože obchody byly zavřené. Občas si lidé objednávali jídlo domů. Začali nakupovat více věcí jako roušky, dezinfekce a třeba i počítače nebo tablety na online výuku.

Myslíte si, že se lišilo, co lidé nakupovali v různých zemích/kontinentech?

Možná odpověď: Pravděpodobně ano. Lidé potrebovali různé věci například na základě toho, jaká opatření byla zavedena.

Uvědomění

35
minut10
minut

Aktivita 1

Promítněte žákům stránku [The New Normal](#) (prezentace strana O3).

Ten vizualizuje, jak se proměnilo nákupní chování na základě vyhledávacích dotazů uživatelů ve vyhledávači Google před, během a po období pandemie koronaviru.

Vzorce chování uživatelů řadí do tří kategorií:

New Normal: Počet vyhledávání se během pandemie změnil a změna se stala novou normou.

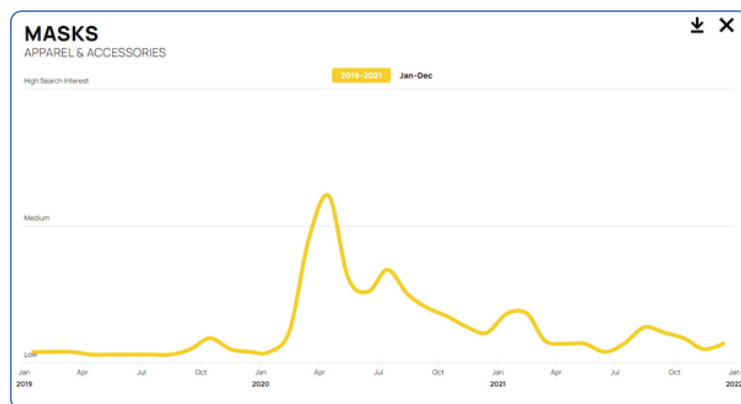
Unusual: Počet vyhledávání se během pandemie změnil, ale poté se vrátil k předešlým hodnotám.

Normal: Počet vyhledávání zůstal pandemií relativně nedotčen.

Projděte s žáky všechny tři kategorie. Zaměřte se na takové výsledky (grafy), které jsou žákům blízké. Doporučujeme představit například graf vyhledávání pojmu „Masks“.



The new normal



Diskutujte s žáky o důvodech zvýšeného zájmu lidí, kteří v období od ledna 2020 do přibližně června 2022 vyhledávali prostřednictvím vyhledávače Google pojem „Mask“. Všimněte si také, jak se zájem propadl po dvou letech, tedy v lednu 2022.

10
minut

Aktivita 02

Promítněte žákům stránku [What are we searching for](#) (prezentace strana 04).

Stránky předkládají analýzu chování lidí jako „den v životě“ v různých částech světa. Zjistíte například, že lidé v České republice vyhledávají termín související s bankou především ráno. Podobně jako lidé ve Finsku, Indii nebo Peru. Nebo, že zájem o pizzu roste v některých státech především v podvečerních nebo večerních hodinách. Viz obrázky níže.



[What Are We Searching For](#)



Promítněte prezentaci na straně 05.

Dokonči větu

Nedokončené věty

Žáci dokončí věty: 1) Věděli jsme, že... 2) Zaujalo nás, že... 3) Překvapilo nás, že... 4) Mýlili jsem se v tom, že... 5) Nerozuměli jsme...

Rozdělte žáky do skupin (každá skupina bude potřebovat jedno zařízení). Žáci si prohlédnou stránku [What are we searching for](#) a na papír zapíší doplnění vět, které vidí v prezentaci na straně 05. Následně mezi sebou v diskuzi sdílí ty, které považují za nejzajímavější.

15
minut

Aktivita 03

Analyzujte data z období pandemie covidu pomocí nástroje Google Trends (prezentace strana 06).

Pracujte s nástrojem [Google Trends](https://trends.google.com/trends/explore): trends.google.com/trends/explore. Nástroj vizualizuje vyhledávání výrazů ve vyhledávači Google. Můžete studovat informace o místě a čase různých vyhledávaných výrazů, včetně jejich srovnání s jinými.

Hledejte postupně výrazy uvedené v seznamu níže. Nezapomeňte upravit období vyhledávání na cca únor 2020 až do prosince 2022 (omezíme časovou řadu na období pandemie koronaviru a přidáme několik měsíců navíc, abychom viděli případnou změnu v dalším vývoji). Zanechte region Česká republika. Čísla vyhledávání jsou relativní (neznáme tedy počet vyhledávání, ale pouze jejich poměr).

Hledejte například: Gumičky, šicí stroj, jízdenka, alkohol, sázení, běžky, běžecké trasy.

Nechte žáky interpretovat, proč zrovna tyto výrazy byly takto vyhledávané.



[Google Trends](#)



Gumičky



Šicí stroj

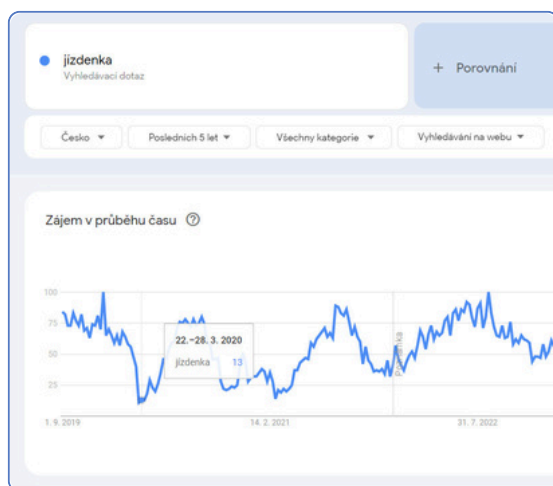
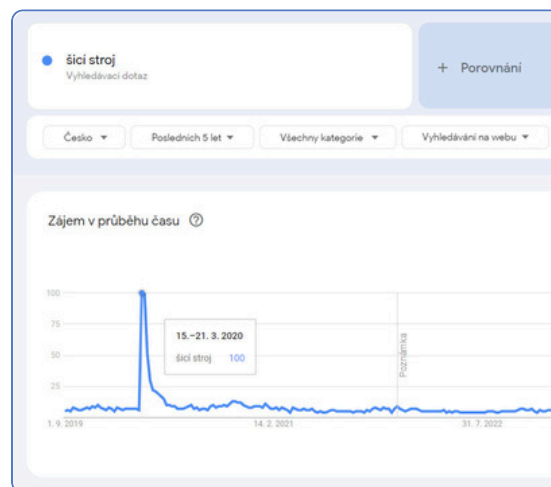
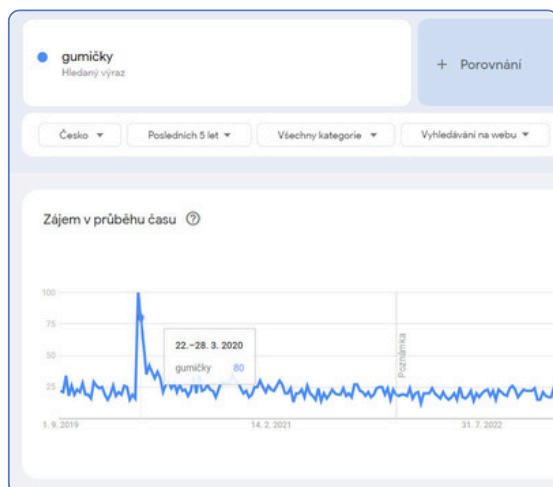


Jízděnka



Lístečková
metoda

Výsledky vyhledávání výrazů „gumičky“ a „šicí stroj“ a „jízděnka“ (prezentace strana 07).



Žáci pracují ve dvojicích a hledají další výrazy spojené s pandemií Covid-19.

Lístečková metoda

Žáci zapisují vyhledávané výrazy na lístečky a poté hromadně lístečky shlukují dle jejich podobnosti.

V následné diskusi společně interpretujte ty nejčastější nebo nejzajímavější lístečky nebo jejich shluky.

Reflexe

5
minutZamysli se

Napadají tě další témata nebo období, která by byla zajímavá interpretovat z velkých dat?

Co si představíš pod pojmem velká data?

Ustálená definice pojmu velká data (Big Data) neexistuje, avšak společné charakteristiky velkých dat zahrnují jejich ohromnou velikost, to jak rychle vznikají, jak jsou různorodá (velikostí a strukturou) a také, že mají nižší kvalitu nebo nejsou tak přesná. Velká data mohou zahrnovat obrázky, videa, audia, texty nebo tzv. digitální stopy ve formě údajů o uživatelském chování. Jsou důsledkem zrychlení a vývoje internetu, přičemž roli hraje také výrazné zlevnění ukládání a zpracování dat. Typicky je jich tolik, že jejich zpracování vyžaduje nové přístupy, například právě metody strojového učení.

Kolik myslíš, že lidé provedou vyhledávání na Google denně?

Dle článku na stránce seo.aj uživatelé zadají cca 99 000 vyhledávacích dotazů každou vteřinu, což znamená zhruba 8,5 miliard vyhledávacích dotazů každý den.

Kde jinde myslíš, že vznikají velká data?

Možné odpovědi: Když interagujeme na sociálních sítích. Při transakcích v online obchodech nebo bankách. Ve veřejné správě a při monitorování dopravy a bezpečnosti. Ve zdravotnictví, při sběru dat z lékařských záznamů a lékařských zařízení. Ve veřejné správě a při monitorování dopravy a bezpečnosti.

K čemu bychom mohli dále využít velká data?

Možné odpovědi: Pro předpovídání počasí (senzory na celém světě měří teplotu apod.) nebo pro předpověď a zmírnění dopadů přírodních katastrof. Zlepšení zemědělských výnosů prostřednictvím analýzy půdních dat, klimatu a používání vody...