



AI dětem

Kurikulum umělé inteligence pro ZŠ a SŠ

Karta

Předpojatost

03

balíček Strojové učení



kurikulum.aidetem.cz/informatika2

Vypracovali: Radek Špáta a Eva Nečasová
Odborní garanti: Pavel Kordík, Tomáš Mlynář
Jazyková korektura: zatím neproběhla
Poslední aktualizace: 06/2024
Verze: 03



[Formulář pro
přípomínky.](#)

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

Předpojatost

Toto by bylo dobré vědět o předpojatosti pro začátek

Představte si, že chcete vytvořit aplikaci umělé inteligence, která rozpoznává na obrázcích boty. Nejprve jí musíte ukázat mnoho obrázků bot, aby na nich aplikace našla podobnosti (například tkaničky, podrážky...). A teď si zkuste představit co nejvíce typů bot. Všechny tyto boty byste aplikaci ukázali. Ale teď pozor. Představili jste si nejen všechny typy bot, ale ze všech úhlů, ve všech barvách, materiálech.... se všemi jednotlivými detaily? Boty jaké se nosili ve starověkém Egyptě nebo v Číně v 5. st.? AI nevnímá věci tak jako my lidé. Nemůže si boty osahat, nedokáže generalizovat tak úspěšně jako my. Proto některé příklady, které byste aplikaci neukázali, by jako boty nemusela rozpoznat. A tomu se říká předpojatost.

Rozšíření definice pro pokročilé

Systémy umělé inteligence se učí na datech, která připravují lidé. Špatně připravená, nevyvážená data nebo jejich nedostatek mohou způsobit, že umělá inteligence bude určitým způsobem předpojatá nebo také zkreslená (v angličtině biased). Předpojatost může být například historická, kdy se dlouho opakované lidské předsudky (rasa, gender...) promítnou také do dat – umělá inteligence bude na vedoucí pozici doporučovat spíše muže, protože v minulosti to tak doopravdy bylo. Nebo reprezentační – bude rozpoznávání tváří správně fungovat v Africe, když jsme natrénovali systém umělé inteligence na obličejích lidí ze střední Evropy?

Pokud má umělá inteligence hrát větší roli v rozhodovacích procesech, je důležité, abychom se na ni byli schopni spolehnout s ohledem na její přesnost a nezájatost. Proto je velmi důležité, abychom AI systémy učili na těch správných datech.

Balíček Strojové učení

Dataset
Strojové učení s učitelem
Předpojatost
Posilované strojové učení



[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

Mimozemská detektivní kancelář III. díl

45
minut

Popis aktivity

V předchozím cvičení jsme trénovali model strojového učení v aplikaci Teachable Machine s daty dvou mimozemských rodin (Fluffů a Earlů). Od modelu jsme očekávali, že bude s velkou mírou pravděpodobnosti správně klasifikovat obrázky nalezců (přidělí je do správné rodiny). Zjistili jsme ale, že pro náš model byly zásadním rozlišovacím atributem uši, proto nalezenec klasifikoval jako člena rodiny Earlů. Svou teorii jsme mohli prověřit tak, že jsme model testovali na obrázcích dalších nalezců. Obrázek nalezenec bez uší byl přiřazován do skupiny Fluffů. Jinými slovy... ačkoli měl náš první nalezenec velmi podobné atributy s nalezcem bez uší (zdálo by se, že patří do stejné rodiny), byli oba nakonec klasifikováni jako členové jiných rodin.

Následující aktivita přímo navazuje na předchozí úkol z karty Strojové učení z učitelem. V praktickém cvičení využijete předem připravený projekt v aplikaci Teachable Machine. Díky dobrodružné expedici na mimozemské planetě se však objeví další dvě rodiny, které do projektu zahrnete a tím rozšíříte trénovací data. Nicméně ani to nepomůže správně přiřadit nalezenec a model bude stejně předpojatý. Změnu konečně přinese až expedice do ne zakázaných končin planety Yemi.

Jak souvisí detektivní kancelář s předpojatostí

Natrénujeme-li model strojového učení na datech, která nejsou jednoznačná a je jich příliš málo (tak jako je to v případě obou rodin), je pro něj velmi těžké rozpoznat (klasifikovat) nová data (nalezenec) správně. Uvidíme, že ani rozšíření datasetu o dvě další rodiny nepomůže. Stále bude atribut uší ten hlavní, díky němuž jsou nalezcem přidělováni do nesprávných rodin. V „realitě“ ale nemusí být uší ten správný atribut. Může to být srst či jiný detail, který jsme ani nezaznamenali. Ale přesto model na jeho základě rozhoduje – je tedy předpojatý.

Informace o aktivitě

Prekoncepty/ročníky, délka lekce

6.–9. ročníky, 45 minut

Klíčová slova

Předpojatost, dataset, strojové učení

Co se žáci učí?

Systémy strojového učení jsou předpojaté.
Předpojatosti lze předcházet testováním modelu a úpravou datasetu.

Proč se to učí?

Cíl pro balíček karet Strojové učení:
Kriticky posoudí rozhodování systémů umělé inteligence.

Jak poznáme, že se to naučili?

Vysvětlí pojem předpojatost.

Aktivita

Žáci vylepšují dataset a testují model strojového učení.

Pomůcky

Pedagog: sady kartiček v rodinami mimozemšťanů, projekční zařízení, prezentaci
Žáci: Počítač, notebook nebo tablet (ale ne s OS Android) pro každého žáka nebo do skupiny.

Výstupy RVP – Informatika

I-9-1-04: zhodnotí, zda jsou v modelu všechna data potřebná k řešení problému; vyhledá chybu v modelu a opraví ji.

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-C-III Datové sady (předpojatost)

Bloomova taxonomie

Porozumění: Chápu, jak nedostatečná nebo špatně reprezentativní data vedou k předpojatosti v modelech AI.
Aplikace: Používají nástroj Teachable Machine pro trénování a testování modelu, identifikují případy předpojatosti v datech.

Evokace

5 minut

Vybav si

V předchozí lekci jsme trénovali model strojového učení v aplikaci Teachable Machine s daty dvou mimozemských rodin (Fluffů a Earlů). Od modelu jsme očekávali, že správně přidělí nalezenec do správné rodiny. Zjistili jsme ale, že pro náš model byly zásadním rozlišovacím atributem uši. Na základě čeho jsme si to ověřili?

Svou teorii jsme prověřili tak, že jsme model testovali na obrázcích dalších nalezců. Obrázek nalezenec bez uší byl jako jediný přiřazován do skupiny Fluffů. Jinými slovy... ačkoli měl náš první nalezenec velmi podobné atributy s nalezcem bez uší (zdálo by se, že patří do stejné rodiny), byli oba nakonec klasifikováni jako členové jiných rodin.

Vysvětlí

Model v aplikaci Teachable Machine jsme trénovali pomocí tzv. strojového učení s učitelem. Dokážeš popsat, jak tento typ strojového učení funguje?

Stroje se mohou, podobně jako lidé, učit z příkladů. Využívají k tomu mimojiné metodu, která se nazývá strojové učení s učitelem. Zjednodušeně to znamená, že nemusíme strojům dávat přesná zadání, nebo chcete-li postup (algoritmus). Místo toho jim ukážeme příklady, ze kterých se sami naučí vykonávat různé úlohy. Lidé jsou učitelé, kteří modelu říkají, co je co – například pomocí štítků, třídění do tříd/kategorií apod.

Zamysli se

Stále nevíme, do jaké rodiny nalezenec patří. Co by nám mohlo pomoci?

Rozšířit dataset o více dat s obrázky mimozemšťanů. Mít více informací o nalezcích a rodinách. Snažit se nalezenec klasifikovat pomocí jiných dat – mimo obrázků například pomocí DNA apod.

Uvědomění

30 minut

Aktivita 1

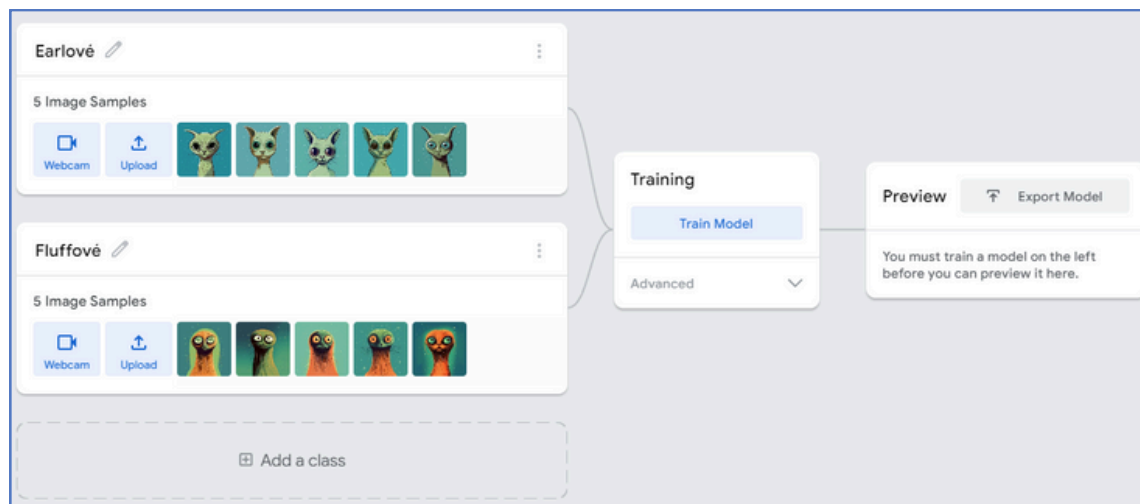
5 minut

Žáci otevřou webovou stránku Teachable Machine: teachablemachine.withgoogle.com

- + Zvolí Get started + Open an existing project from a file.
- + Nasdílejte jim předpřipravený projekt „2-rodiny.tm“ – bit.ly/mimozemstani03.
- + Žáci si také mohou naskenovat nebo QR kód v prezentaci na straně 02.
- + Po nahrání uvidí projekt se dvěma třídami – Fluffové a Earlové.



2-rodiny.tm



Příběh pokračuje... (v prezentaci na straně 03)

Místní média publikují řadu informací z expedice do neprozkoumané oblasti planety Yemi. Zdá se, že dokonce objevili nové rodiny mimozemšťanů! Jeden z našich kolegů v detektivní kanceláři je současně spolupracovníkem expedice. Díky němu získáváme exklusivně přístup k novým důkazům!

15 minut



nová data

Sdílejte s žáky nová data (odkaz a QR kód v prezentaci na straně 04).

Instruuje žáky, ať se seznámí s novými daty, která poskytla expedice. Je možné využít některá data k vylepšení naší modelu a pokusit se vyřešit hádanku s přidělením nalezence do správné rodiny? Data najdete na bit.ly/vzorky-expedice.

Žáci doplní dataset v Teachable Machine. Vytvoří nové třídy, pojmenují je a nahrají do nich obrázky. Sledujte, zda žáci při nahrávání obrázků volí vhodné obrázky (tj. mimozemšťany) či zda při tvorbě nové třídy vkládají i obrázky, které do třídy logicky nepatří (rostliny, jiní živočichové, ...). Nechte žáky vytvořit dataset jakkoliv. Po vytrénování modelu a vyzkoušení klasifikace nalezence můžete s chybou pracovat a analyzovat, proč neuronová síť nalezence stále přiřazuje do „nehodné“ třídy. Tzn. novou třídu vylepšit a nevhodné obrázky ze třídy odstranit apod.

10 minut



nalezenci

Žáci otestují natrénovaný model pomocí původních obrázků nalezence.

Pokud nemáte schované obrázky nalezenců z minulé lekce, jsou zde: bit.ly/nalezenci

Nyní většinou nalezenci spadají do rodiny Kittinové.

A co tváře žáků? 😊

Ukažte žákům v prezentaci vysvětlení předpojatosti na straně 08.

Povídejte si tom, kdy jsou modely strojového učení předpojaté. Napadají žáky podobné příklady jako s botami? Podobné například mohou být: ovoce a zelenina, tváře lidí v Evropě vs. tváře lidí v Asii a podobně.

Reflexe

5
minutSdílej
a argumentuj**Do jaké třídy které nalezenec tvůj model přiřadil a proč?**

Odpověď bude pravděpodobně do rodiny Kittinů. Nalezenec bez uší byl stále určován jako Fluff.

Myslíš, že v tom znovu hrál roli výrazný atribut uší?

Dle našeho testování i s novými daty byl atribut uší stále rozhodující.

Model fungoval technicky správně (tj. snažil se rozpoznávat obrázky), ale přesto nejsme schopni říci, zda přiřadil nalezenec do správné rodiny. Jak je to možné?

Nemáme dostatek dat a nevíme, jak to ve skutečnosti bylo. Proto ani nemůžeme potvrdit, že model nalezenec přiřadil správně nebo špatně.

Na základě naší zkušenosti s trénováním a testováním modelu strojového učení – myslíš si, že můžeme (je správné) systémům AI důvěřovat na 100 %? Proč?

Umělá inteligence si vytváří vlastní interpretaci na základě dat, která jí předložíme. Je možné, že lidé připraví taková data nesprávně, případně málo otestují model a ten pak nemusí fungovat správně. Proto je velmi důležité kriticky uvažovat na rozhodnutích, která jsou automatizována pomocí AI.

Jak poznáme, že je systém umělé inteligence (model) předpojatý?

Na základě důsledného testování.

Jak si myslíš, že by to nakonec s nalezcem mohlo být? Do které rodiny dle tebe „doopravdy“ patřil?

Vyústění příběhu: Během expedice nebyla rodina malého nalezenec vůbec objevena, a proto tato data v datasetu úplně chyběla. Proto model nefungoval správně a přiřazoval nalezenec do nesprávných rodin. Později se sama rodina malého nalezenec přihlásila, protože zachytila signál se zprávou, aby se všechny rodiny na planetě Yemi přihlásily. Nalezenec tedy nakonec svou rodinu našel!

Ukažte žákům stránku 09 v prezentaci. 🍷

Navazující lekce

V této lekci žáci zjistili, co je předpojatost.

Příběh mimozemské detektivní kanceláře končí. Navazuje poslední lekce z balíčku Strojové učení, která žákům ukáže další typ strojového učení – posilovaného.