co říkáte na tento projekt k hacktonu?

Projekt „Predikce atraktivity Olomouckého kraje pro nové firmy a startupy“ si klade za cíl využít machine learning k identifikaci nejvhodnějších oblastí v Olomouckém kraji pro zakládání nových firem a startupů. V rámci tohoto projektu se zaměříme na analýzu různých faktorů, které mohou přispět k úspěchu nových podniků, jako je infrastruktura, blízkost univerzit, dostupnost financí a počet startupů v okolí.

Pro realizaci projektu bude nejprve nutné shromáždit relevantní data. Budeme sbírat demografická data, která zahrnují obyvatelstvo, vzdělanost, migraci talentů, průměrný věk obyvatel a kvalifikaci pracovní síly. Dále se zaměříme na ekonomická data, jako jsou počty nově vzniklých a zaniklých firem, míra nezaměstnanosti, průměrné mzdy, dostupnost financí a objem investic. Důležitým faktorem bude také analýza infrastruktury, kde budeme posuzovat blízkost výzkumných institucí, univerzit, technologických parků a dopravní dostupnost. Zahrneme i data o startupovém ekosystému, která zahrnují počty startupů, inkubátory, dostupnost coworkingů a úspěšnost startupů. Historická data o úspěšnosti firem nám poskytnou přehled o tom, které firmy v regionu vznikly a jak se vyvíjely.

Pro modelování použijeme machine learning, přičemž cílovou proměnnou bude úspěšnost nových firem, například míra přežití po třech letech, růst zisku a počet zaměstnanců. Vstupní proměnné budou zahrnovat faktory z výše uvedených datasetů. Vytvoříme klasifikační modely, například Random Forest nebo Gradient Boosting, které nám umožní predikovat, zda nová firma založená v určité lokalitě bude úspěšná. Dále implementujeme regresní modely, které budou sloužit k predikci míry růstu firmy, například v oblasti tržeb a počtu zaměstnanců.

Výsledky našeho modelování budeme vizualizovat pomocí interaktivní mapy, která zobrazí predikci atraktivity jednotlivých oblastí Olomouckého kraje pro startupy a nové firmy. Tato mapa bude obsahovat různé vrstvy, které ukáží například top 5 nejvhodnějších lokalit pro nové startupy a faktory ovlivňující jejich atraktivitu, jako je blízkost univerzit, dostupnost financí a dopravní infrastruktura.

Na závěr projektu identifikujeme konkrétní oblasti Olomouckého kraje, které mají největší potenciál pro rozvoj nových firem v inovačních oblastech. Na základě našich zjištění navrhneme doporučení pro politiky a investice, které mohou zvýšit atraktivitu dalších regionů, například zlepšením infrastruktury a podporou startupového ekosystému.

Tento projekt přinese významné přínosy tím, že vytvoří nástroj, který pomůže investorům, podnikatelům a státní správě rozhodovat o tom, kde investovat do rozvoje nových firem a inovačních projektů. Zároveň identifikujeme klíčové faktory, které přispívají k úspěchu startupů, a navrhneme, jak zlepšit podmínky pro jejich růst, což v konečném důsledku přispěje k rozvoji Olomouckého kraje.

**Dataset: Počet a typy výzkumných institucí a univerzit**

* **Nápad**: Identifikovat, kolik výzkumných institucí, univerzit a fakult zaměřených na inovace v Olomouckém kraji působí. Porovnat s jinými kraji a analyzovat, zda v některých oblastech vědy a výzkumu existuje potenciál růstu (např. biotechnologie, medicína, IT). Vizualizace může ukázat, kde má kraj silné a slabé stránky.
* **Možné zdroje dat**: Český statistický úřad, Ministerstvo školství, data vysokých škol a výzkumných ústavů.

**2. Dataset: Počet patentů a inovací**

* **Nápad**: Analyzovat počet podaných patentů v Olomouckém kraji a porovnat s jinými kraji. Identifikovat oblasti s vysokou patentovou aktivitou a oblasti, kde by mohla být podpora inovací a patentování zvýšena. Také sledovat, jaké firmy a výzkumné instituce mají nejvyšší míru inovací.
* **Možné zdroje dat**: Databáze patentů, Český úřad průmyslového vlastnictví.

**3. Dataset: Financování vědy a výzkumu**

* **Nápad**: Zjistit, kolik financí plyne do vědy a výzkumu v Olomouckém kraji (např. z grantů, veřejného sektoru, evropských fondů). Porovnat to s jinými kraji a zjistit, zda existuje potenciál pro zlepšení financování v určitých oborech. Identifikovat investiční mezery a navrhnout strategie na jejich řešení.
* **Možné zdroje dat**: Data z Národního informačního systému pro výzkum a vývoj (IS VaVaI), data Evropských strukturálních fondů.

**4. Dataset: Míra zaměstnanosti v inovačních odvětvích**

* **Nápad**: Prozkoumat zaměstnanost v sektorech souvisejících s vědou, výzkumem a inovacemi (IT, biotechnologie, strojírenství, vývoj nových materiálů apod.). Porovnat míru zaměstnanosti v těchto odvětvích s ostatními kraji a identifikovat oblasti, kde by bylo vhodné podpořit rozvoj pracovních míst.
* **Možné zdroje dat**: Úřad práce, Český statistický úřad.

**5. Dataset: Demografie a vzdělanost obyvatel**

* **Nápad**: Prozkoumat demografická data o obyvatelích Olomouckého kraje, jejich vzdělání, podíl vysokoškoláků v oborech jako technické vědy, IT, přírodní vědy. Analyzovat, jaká je migrace talentů – zda vzdělaní lidé z kraje odcházejí nebo zůstávají, a jak přilákat nové odborníky.
* **Možné zdroje dat**: Český statistický úřad, data univerzit, MPSV.

**6. Dataset: Startupový ekosystém a investice**

* **Nápad**: Zmapovat existenci startupů v Olomouckém kraji, jejich zaměření a úspěšnost. Analyzovat, jaké jsou investiční možnosti pro startupy v kraji a jaký je přístup k rizikovému kapitálu. Na základě dat navrhnout konkrétní oblasti, kde by mohl kraj startupům více pomoci (např. technologické inkubátory, investiční fondy).
* **Možné zdroje dat**: Startupové databáze (např. Crunchbase, AngelList), CzechInvest.

**7. Dataset: Kvalita života a infrastruktura pro vědu**

* **Nápad**: Prozkoumat kvalitu života v Olomouckém kraji, dostupnost bydlení, infrastrukturu (dopravní spojení, dostupnost coworkingových center, vědeckých parků). Identifikovat, jak kvalita infrastruktury ovlivňuje atraktivitu kraje pro inovátory a výzkumníky.
* **Možné zdroje dat**: Český statistický úřad, data měst a obcí, dopravní infrastruktura.

Tyto datasety a nápady mohou sloužit jako základ pro návrh projektu zaměřeného na zvýšení atraktivity Olomouckého kraje, identifikaci jeho silných stránek a příležitostí pro zlepšení v oblasti vědy, výzkumu a inovací.

**Cíl projektu:**

Využít machine learning pro predikci, které oblasti Olomouckého kraje jsou nejvhodnější pro zakládání nových firem a startupů, a identifikovat, jaké faktory (např. infrastruktura, blízkost univerzit, dostupnost financí, počet startupů v okolí) nejvíce přispívají k úspěchu nových podniků v inovačních oblastech.

**Popis řešení:**

1. **Sběr dat:**
   * **Demografická data**: Obyvatelstvo, vzdělanost, migrace talentů, průměrný věk obyvatel, kvalifikace pracovní síly.
   * **Ekonomická data**: Počet nově vzniklých a zaniklých firem, míra nezaměstnanosti, průměrná mzda, dostupnost financí, objem investic.
   * **Infrastruktura**: Blízkost výzkumných institucí, univerzit, technologických parků, dopravní dostupnost.
   * **Startupový ekosystém**: Počet startupů, inkubátory, dostupnost coworkingů, úspěšnost startupů.
   * **Historická data o úspěšnosti firem**: Data o firmách, které v kraji vznikly, a informace o jejich růstu nebo krachu.
2. **Modelování pomocí machine learningu:**
   * **Cílová proměnná**: Úspěšnost nových firem (např. míra přežití po 3 letech, růst zisku, počet zaměstnanců).
   * **Vstupní proměnné**: Různé faktory z výše zmíněných datasetů (demografie, ekonomická data, infrastruktura, apod.).
   * **Modely**:
     + **Klasifikační model** (např. **Random Forest**, **Gradient Boosting**): Predikce, zda nová firma založená v určité lokalitě bude úspěšná na základě historických dat a místních podmínek.
     + **Regresní model**: Predikce míry růstu firmy (tržeb, zaměstnanců) v dané lokalitě.
3. **Vizualizace výsledků:**
   * Interaktivní mapa, která bude zobrazovat predikci atraktivity jednotlivých oblastí Olomouckého kraje pro startupy a nové firmy.
   * Mapa by mohla zahrnovat různé vrstvy, které ukazují:
     + **Top 5 nejvhodnějších lokalit** pro nové startupy.
     + **Faktory ovlivňující atraktivitu**: například blízkost univerzit, dostupnost financí, dopravní infrastruktura.
4. **Závěry a doporučení:**
   * Identifikovat konkrétní oblasti Olomouckého kraje, kde je největší potenciál pro rozvoj nových firem v inovačních oblastech.
   * Doporučit politiky a investice, které mohou zvýšit atraktivitu ostatních oblastí – např. zlepšení infrastruktury, podpora startupového ekosystému.

**Datasety, které mohou být použity:**

* **Data o firmách a startupovém ekosystému** (z CzechInvest, Crunchbase, AngelList).
* **Demografická a ekonomická data** (Český statistický úřad, data o zaměstnanosti a příjmech).
* **Data o vzdělávacích a výzkumných institucích** (Ministerstvo školství, univerzity, výzkumné instituce).
* **Infrastruktura a dopravní dostupnost** (OpenStreetMap, data měst a krajů).

**Přínosy:**

* Vytvoření nástroje, který může pomoci investorům, podnikatelům i státní správě rozhodovat o tom, kde investovat do rozvoje nových firem a inovačních projektů.
* Identifikace, které faktory nejvíce přispívají k úspěchu startupů, a návrhy, jak zlepšit podmínky pro jejich růst.

**1. Titulní strana**

* Název projektu: **"Predikce atraktivity Olomouckého kraje pro nové firmy a startupy"**
* Jména týmu členů
* Hackathon: **"Atraktivita Olomouckého kraje v oblasti vědy, výzkumu a inovací"**
* Datum

**2. Úvod a kontext problému**

* Stručně vysvětlete, proč je důležité podporovat startupový ekosystém a podnikání v Olomouckém kraji.
* Ukažte hlavní výzvy regionu, např. nižší míra zakládání firem, menší počet inovačních projektů oproti jiným krajům ČR.
* Jak by mohly startupy a nové firmy pomoci ekonomickému růstu kraje.

**3. Cíl projektu**

* Vysvětlete, co chcete s projektem dosáhnout:
  + Identifikovat nejvhodnější oblasti v Olomouckém kraji pro nové startupy a firmy.
  + Využít machine learning pro predikci úspěšnosti firem na základě různých faktorů (infrastruktura, demografie, ekonomika atd.).
  + Navrhnout, jaké kroky mohou zlepšit atraktivitu regionu.

**4. Datasety**

* Přehled použitých datasetů:
  + Demografická data (obyvatelstvo, vzdělanost, migrace).
  + Ekonomická data (nezaměstnanost, průměrná mzda, počet firem).
  + Data o infrastruktuře (blízkost univerzit, výzkumných center, doprava).
  + Historická data o firmách (růst, přežití startupů).
* Zdůrazněte důležitost těchto dat pro modelování a predikci.

**5. Metodologie: Machine Learning**

* Krátký úvod do použitých technologií a algoritmů.
  + Jaký typ modelu používáte (např. klasifikace pro predikci úspěchu firem).
  + Proč byl vybrán konkrétní machine learning model (např. Random Forest, Gradient Boosting).
* Stručně vysvětlete, jak probíhá trénování modelu a jaké vstupní proměnné používáte.

**6. Analýza faktorů atraktivity**

* Prezentujte výsledky analýzy, která identifikuje klíčové faktory pro úspěch startupů v kraji.
  + Např. blízkost univerzit, nízká nezaměstnanost, dopravní dostupnost, přístup k financím.
* Ukázky grafů nebo tabulek s nejdůležitějšími faktory.

**7. Výstup modelu: Predikce atraktivity**

* Vizualizace predikce (např. heatmapa Olomouckého kraje s vyznačením atraktivních oblastí).
* Ukázka top 5 lokalit, kde by nové firmy mohly mít nejvyšší úspěšnost.
* Vizualizace míst s potenciálem pro růst, ale s nutností zlepšení podmínek.

**8. Návrh doporučení**

* Jak na základě výsledků zlepšit atraktivitu dalších oblastí v kraji.
* Např. zlepšení infrastruktury, podpora startupů, investice do výzkumu a vývoje, propojení s univerzitami.

**9. Možné využití v praxi**

* Jak mohou výsledky projektu využít:
  + Investoři (kde zakládat nové firmy, kde investovat).
  + Místní vláda (kde soustředit podporu a zdroje pro rozvoj podnikání).
  + Startupové inkubátory (výběr lokalit pro nové projekty).
* Přínosy pro ekonomiku a inovační ekosystém kraje.

**10. Závěr a další kroky**

* Shrnutí hlavních výstupů.
* Jak by mohl být model dále vylepšen (např. přidání dalších datových zdrojů, jemnější predikce).
* Výzva k dalšímu rozvoji projektu (možné partnery, zájemce o spolupráci).