

Fundamentele algoritmice și etica în IA: de la teorie la practică

Set de instrumente pentru sesiuni sincrone

UC3 | Algoritmi și limitările acestora
Studiu de caz

Studiu de caz



Studiu de caz

Descriere pentru formatori

	Descriere
Descrierea sarcinii	Students will analyze a fictional case (see next slide) involving a financial technology (FinTech) company that used an algorithmic system to assess creditworthiness for loan approvals. The company aimed to provide quick and efficient loan services to a broader demographic. However, it was later discovered that the algorithm consistently favored applicants from certain income brackets and racial backgrounds, resulting in discriminatory loan practices. Students must investigate the root causes of this issue, evaluate the consequences on affected communities, and suggest effective mitigation strategies.
Descrierea modului de realizare a sarcinii	Lucru în grup Grupuri de 3-4 persoane. Cercetare și analiză de fond • Revizuiți materialele cursului și literatura de specialitate privind prejudecățile algoritmice și echitatea. • Înțelegeți factorii specifici care au contribuit la rezultatele discriminatorii ale FinTech Innovations, inclusiv prejudecățile legate de date, caracteristicile utilizate și absența indicatorilor de echitate. Identificarea cauzelor profunde • Analizați modul în care datele și caracteristicile de formare ale FinTech Innovations ar fi putut integra prejudecăți sociale. • Investigați absența indicatorilor de echitate și lipsa contribuției părților interesate. Recomandări de atenuare Elaborați o serie de recomandări pentru a aborda prejudecățile identificate, inclusiv diversificarea datelor, noi proceduri de validare, orientări etice și audit continuu. Redactare și prezentare Organizați concluziile într-un raport structurat, oferind un rezumat clar, o analiză a cauzelor principale și strategii detaliate de atenuare. Includeți imagini, grafice sau tabele, acolo unde este relevant, pentru a vă susține argumentele.
Timpul estimat pentru realizarea sarcinii	30-60 minute.
Sugestii de surse pentru realizarea sarcinii	Materiale de curs, baze de date academice, sisteme GAI (de exemplu, ChatGPT, Copilot etc.) și căutări pe internet.
Descrierea detaliată a modului de realizare a sarcinii	Scrieți un raport sumar de 500-750 de cuvinte.
Informații privind termenul limită pentru predarea sarcinii	În timpul sesiunii sincrone.
Informații de contact sau cum să clarificați nelămuririle	Profesorul trebuie să furnizeze o formă de contact (poate fi o adresă de e-mail, un număr de telefon...).

Studiu de caz

Caz fictiv pentru studenți

Industria FinTech este cunoscută pentru abordarea sa inovatoare în furnizarea de servicii financiare rapide și eficiente, adesea prin utilizarea sistemelor algoritmice pentru evaluarea bonității și aprobarea împrumuturilor. Acest caz fictiv se învârtă în jurul unei companii FinTech care a utilizat un algoritm avansat de evaluare a bonității pentru a evalua solicitanții și a simplifica procesul de aprobare a împrumuturilor. Misiunea companiei era de a democratiza accesul la resurse financiare și de a sprijini clienții care sunt adesea defavorizați de băncile tradiționale. Cu toate acestea, algoritmul său a dus în mod neașteptat la practici discriminatorii de acordare a creditelor, favorizând în mod sistematic solicitanții din anumite categorii de venituri și din anumite medii etnice.

Contextul companiei

Compania fictivă FinTech Innovations a fost înființată pentru a revoluționa sectorul bancar tradițional, oferind un sistem de aprobare a împrumuturilor complet online și fără cusur. Acest sistem utiliza un algoritm propriu pentru a evalua riscul de credit pe baza diverselor factori, precum veniturile, istoricul profesional, locația geografică și nivelul de studii. FinTech Innovations își propunea să pătrundă pe piețele ignorate de băncile tradiționale, în special în rândul tinerilor, al minorităților și al persoanelor cu surse de venit neconvenționale.

Identificarea problemei

În ciuda obiectivelor declarate de incluziune, algoritmul de evaluare a bonității al FinTech Innovations favoriza în mod constant solicitanții cu venituri mai mari și din anumite grupuri etnice. Acest lucru a dus la rate disproporționate de mari de respingere a împrumuturilor pentru solicitanții cu venituri mai mici și cei aparținând minorităților etnice. Discrepanțele în ratele de aprobare a împrumuturilor între grupurile demografice au stârnit îngrijorarea grupurilor de apărare a drepturilor clienților, ceea ce a dus la un audit extern care a confirmat comportamentul discriminatoriu al algoritmului.

Analiza cauzelor fundamentale

Analiza a relevat mai multe cauze fundamentale ale prejudecății. În primul rând, datele istorice utilizate pentru antrenarea algoritmului reflectau disparitățile sociale existente, încorporând prejudecăți sistematice în model. De exemplu, acesta utiliza locația geografică ca caracteristică, care era puternic corelată cu venitul și rasa. Algoritmul se baza în mare măsură pe date tradiționale privind ocuparea forței de muncă, neglijând persoanele cu surse de venit neconvenționale, cum ar fi freelancerii sau lucrătorii ocazionali. În cele din urmă, lipsa indicatorilor de echitate în timpul validării modelului a permis ca aceste prejudecăți să persiste fără a fi detectate.

Evaluarea impactului

Practicile discriminatorii ale FinTech Innovations au avut repercusiuni sociale și financiare semnificative pentru grupurile afectate. Ratele de respingere a creditelor pentru minorități și solicitanții cu venituri mici au fost aproape duble față de cele ale solicitanților cu venituri mai mari, proveniți din medii privilegiate. Această excludere a dus la pierderea oportunităților de creștere financiară și la consolidarea inegalităților existente. Compania s-a confruntat cu prejudicii de imagine, controlul autorităților de reglementare și provocări juridice.

Studiu de caz

Caz fictiv pentru studenți

Strategii de atenuare

Pentru a aborda aceste probleme, FinTech Innovations a implementat mai multe măsuri corective:

- **Diversificarea datelor**

- Au extins setul de date de instruire pentru a include candidați diversi, provenind din diferite niveluri de venit, rase și tipuri de angajare.

- **Indicatori de echitate**

- Au incorporat indicatori de echitate pentru a se asigura că predicțiile algoritmului sunt imparțiale și reprezentative.

- **Măsuri de transparență**

- Compania s-a angajat să împărtășească mai multe informații despre modul în care algoritmul său evaluează bonitatea.

- **Linii directoare etice**

- Noile linii directoare impun monitorizarea continuă, evaluarea impactului și implicarea părților interesate.

Concluzii și recomandări

Acest caz ilustrează consecințele neintenționate ale luării deciziilor algoritmice în domeniul FinTech. El subliniază importanța utilizării unor date de formare diverse, a validării algoritmilor cu indicatori de echitate și a menținerii transparenței pentru a reduce prejudecățile și discriminarea. Standardele etice și cadrele de reglementare la nivel sectorial sunt esențiale pentru promovarea unor practici FinTech echitabile și incluzive.

MULȚUMESC

Număr proiect | 2022-1-ES01-KA220-HED-000085257



Sprijinul Comisiei Europene pentru realizarea acestei publicații nu constituie o aprobare a conținutului, care reflectă exclusiv opiniile autorilor, iar Comisia nu poate fi trasă la răspundere pentru utilizarea informațiilor conținute în aceasta.

