

ds

**dnevi
slovenske
informatike**

9. in 10. maj 2023

Kongresni center Bernardin, Portorož

30. konferenca Dnevi slovenske informatike
Soustvarjamo digitrajno Slovenijo

Uvajanje umetne inteligence (AI) v Finančni upravi Republike Slovenije (FURS)

Uvajanje prediktivne analitike v FURS

Piramida IT storitev v sistemu za podporo odločanju (BI) v okolju FURS

Podatkovna osnova je centralno podatkovno skladišče SAP BW, kamor se stekajo podatki iz raznolikih transakcijskih sistemov FURS kot tudi zunanjih institucij

Sestavni deli BI okolja v FURS so :

- podatkovni strežnik SAP Hana z operacijskim sistemom Linux
- SAP BW je aplikativno okolje za nalaganje, čiščenje in obdelave podatkov
- SAP Bobj je okolje za pripravo standardnih povpraševanj za uporabnike, kot tudi orodje za ad hoc povpraševanja za napredne uporabnike (izven IT)
- SAP Fraud Management je okolje za pripravo seznamov tveganih davčnih zavezancev glede na poslovna pravila, ki jih določijo uporabniki s svojim strokovnim znanjem
- SAP Data Intelligence je okolje za razvoj modelov prediktivne analitike (PA) za določanje rizičnih davčnih zavezancev (DZ)

Poslovni procesi uvajanja AI v FURS

so naslednji :

- nadzor zahtevkov tujcev za vračilo DDV, bil je prvi proces, v katerega smo vključili AI
- kontrola prejetih obračunov davka, ki je zelo kompleksen poslovni proces, ki je že vključeval poslovna pravila iz izkušenj poslovnih uporabnikov, ki so določala uvrščanje obračunov davka v proces kontrole
- izdelava seznama davčnih zavezancev za inšpekcijski nadzor davka od dohodka pravnih oseb.
Priprava seznama je bil pilotski projekt s software-om KNIME, ki je zelo primeren za prvi stik s prediktivno analitiko.
Pomoč tujih strokovnjakov je bila dobrodošla, saj smo spoznali pomen priprave kvalitetnih podatkov za razvoj modelov prediktivne analitike.
- optimizacija SQL kode v razvoju povpraševanj v SAP BW preko okolja ChatGPT v uradu za informatiko

Cilji uvajanja AI v FURS

- kvalitetnejši izbor davčnih zavezancev za nadzor, manjša je verjetnost, da nadziramo davčne zavezance, ki korektno plačujejo davke
- izvajanje nadzora je povsem ne subjektivno, saj izbor davčnih zavezancev izvaja model PA
- z enakim številom zaposlenih doseči večje učinke v procesih nadzora, večja produktivnost zaposlenih
- usmerjenost v bolj celovite nadzore
- boljše poznavanje podatkov, ki jih uporabljajo modeli PA
- uporaba najmodernejših tehnologij dviguje ugled FURS doma in v tujini
- pomoč pri razvoju programskih rešitev v IT

Značilnosti razvoja modelov prediktivne analitike

Subjekti razvoja in njihove vloge :

- poslovni uporabniki poslovnih procesov, v katerih se uvaja AI
 - Poznajo poslovne procese
 - Skrbniki poslovnih procesov odločajo o spremembah izvajanja poslovnih procesov
 - Povezava s terenom s kontrolorji (izobraževanja, pojasnila ob uvajanju AI, feedback)
- urad za informatiko (delavci službe za poslovno analitiko)
 - Poznajo podatkovne strukture, ki so podlaga za razvoj modelov
 - Projektno vodenje
 - Pojasnjevanje rezultatov PA poslovnim uporabnikom in ožjemu vodstvu
 - Izdelava poročil BI iz področja uvajanja PA
- zunanji izvajalec
 - Razvoj modelov
 - Testiranje modelov
 - Presentacija rezultatov
- ožje vodstvo FURS
 - Odobritev sprememb poslovnih procesov

Značilnosti razvoja modelov prediktivne analitike (PA)

Faze razvoja :

- določitev podatkov za modele PA (obračuni davka, podatki davčnega zavezanca in drugi podatki v zvezi z obdavčitvijo zavezancev)
 - Učna množica (leto 2021)
 - Testna množica (leto 2022)
 - Pomembna je kvaliteta podatkov, ki določa rezultate modelov
- razvoj različnih modelov PA v SAP Data Intelligence
- dopolnjevanje modelov, ki jih razvije okolje SAP Data Intelligence
- določanje kriterijev uspešnosti
- rezultati modelov iz podatkov testne množice se kažejo v porazdelitvi ocene tveganosti med štirimi možnimi izidi
 - Resnično pozitivni primeri (TP)
 - Resnično negativni primeri (TN)
 - Napačno pozitivni primeri (FP)
 - Napačno negativni primeri (FN)
- Izbira med najboljšimi modeli
- Izbor modela
- Uvedba modela v produkcijo v poslovni proces

Prediktivna analitika v kontroli obračunov davka

Kontrola obračunov davka pred uvedbo prediktivne analitike

- obstajala so poslovna pravila, ki so določala ali bo obračun davka uvrščen na namizje kontrolorja
- poslovna pravila so rezultat dolgoletnih izkušenj delavcev FURS v procesu kontrole
- števila obračunov uvrščenih na namizje ni možno enostavno prilagajati

Značilnosti uvedbe prediktivne analitike

- uvajanje modelov PA je postopno, ni takojšnje odprave poslovnih pravil
- poslovna pravila se razdelijo na tim. fiksna, druga poslovna pravila in informacije (nekatera poslovna pravila imajo tudi preventivni značaj)
- prejeti obračuni davka se razdelijo na 3 podmnožice
 - Prejeti obračun davka izpolnjuje fiksna pravila : ne glede na izid PA se uvrsti na namizje kontrolorja
 - Prejeti obračun davka izpolnjuje druga poslovna pravila in je izid tveganost PA večje od določene meje (meja 1)
 - Prejeti obračun ne izpolnjuje poslovnih pravil, lahko se izpišejo informacije in je izid tveganosti PA večji od odločene meje (meja 2)
- s spreminjanjem mej (ki je enostaven postopek) lahko uravnavamo razmerje med drugo in tretjo podmnožico in tudi določamo skupno število obračunov davka na namizje na mesec (kvartal)

Prediktivna analitika v kontroli obračunov davka

Operativni nadzor izvajanje kontrole obračunov po uvedbi prediktivne analitike

- razvili smo posebna poročila, ki dnevno kažejo porazdelitev prejetih obračunov davka po treh podmnožicah
- razvili smo poročila, ki povezujejo učinke kontrole z obračuni davka

S rezultati poročil krmilimo porazdelitve množic in se usmerjamo v tiste obračune, pri katerih je večja verjetnost za rizičnost obračuna.

Razvoj novih modelov prediktivne analitike

Na podlagi rezultatov izidov tveganosti smo namesto enega modela za izračun tveganosti prejetih obračunov davka razvili dva modela, enega za mesečne obračune in drugega za kvartalne obračune.

Vsak zase dajeta boljše rezultate, namesto dveh mej za določitev tveganosti imamo sedaj štiri meje. S tem smo še bolj fleksibilni, v katero populacijo obračunov se bolj usmerjamo.

Tehnološka shema produkcije PA za obračune davka

Tehnološki proces izračuna in uporabe izida tveganosti PA

- eDavki (sistem, ki sprejema obračune davka od davčnih zavezancev)
- SAP ERP , ki podatke iz eDavkov sprejema in jih še pred zapisom v transakcijske tabele obračunov davka zapiše v posebno tabelo prejetih obračunov davka za prenos v okolje SAP Data Intelligence
- SAP BW, ki dnevno pripravlja ažurne podatke o davčnih zavezancih in jih pošlje v SAP ERP
- v SAP ERP se združijo podatki iz SAP BW in podatki prejetega obračuna davka
- iz SAP ERP se pošljejo v okolje SAP Data Intelligence, kjer se za vsak prejeti obračun izračuna koeficient tveganosti obračuna
- iz SAP Data Intelligence se zapisi pošljejo v SAP ERP in dodajo v transakcijske tabele obračunov davka
- v SAP ERP se glede na izpolnjevanje fiksnih poslovnih pravil in vrednosti koeficienta tveganosti glede na mejo 1 in mejo 2 izvede prenos obračuna na namizje kontrolorja

Vse obdelave potekajo v realnem času, do sedaj nismo beležili zastojev pretokov podatkov.

Pilotski projekt KNIME

Pilotski projekt je bil izveden na področju priprave seznama davčnih zavezancev za inšpekcijski pregled na področju davka od dohodka pravnih oseb s pomočjo modela prediktivne analitike

- „freeware“ programski paket KNIME je bil uporabljen po nasvetu strokovnjakov IMF (možen je sicer tudi nakup bolj sofisticirane različice)
- deluje na osebem računalniku
- osnovni model je razvil strokovnjak IMF s podatki , ki smo jih pripravili v IT FURS
- v Uradu za informatiko se je za razvoj modelov usposobil en delavec, modele lahko sedaj v tem okolju razvijamo samostojno
- s tujimi strokovnjaki smo spoznali pomembnost priprave ustreznih podatkov za model kakor tudi obdelavo teh podatkov, da so primernejši za uporabo modelov (npr. prihodke porazdelimo v nekaj razredov, npr. od 1 do10)
- spoznali smo, da lahko cilji določajo, kako se bodo modeli obnašali (npr. uvedba določenih mej za podatke)
- okolje KNIME je zelo primerno za pilotske projekte , saj hitro daje rezultate
- orodje je težko vključiti v obstoječe informacijske sisteme

Primer uporabe rezultatov AI v Uradu za informatiko

Uporaben primer optimizacije SQL programske kode za izvedbo povpraševanja v SAP BW Hana

- v Uradu za informatiko smo uporabili okolje ChatGPT za optimizacijo SQL kode za kompleksno povpraševanje po podatkih iz SAP Hana BW.