

Programm finantsaruannete analüüsimiseks

Karl Artur Lokk

Anna-Liisa Jõgi

Meie projekt keskendub programmi loomisele, mis suudaks automaatselt analüüsida e-Äriregistrist saadud ettevõtte aastaaruandeid. Praegu puuduvad automatiseeritud tööriistad, mis võimaldaksid võrrelda ettevõtte ajaloolist teavet. Kuna aastaaruannetel puudub kindle struktuur siis peaks programm töötama nii struktureeritud tabelite kui ka struktureerimata kirjelduste väljadega. Eesmärk on arendada tööriist, mis hoiab kokku aega ja pakuks ülevaadet kuidas ettevõtte aastate jooksul muutunud on.

Eesmärgid:

1. Luua programm, mis suudab ettevõtte aastaaruannetest lugeda olulist äriteavet.
Programm peaks suutma tuvastada erineva ülesehitusega aruannetest kõige olulisemad parameetrid (tulu, käive, võlad, muutused juhtimises, äriilgis jne) ning võrrelda neid eelmiste aastatega.
2. PDF-aruande automaatne koostamine
Aruandes tuuakse esile muutused finantsnäitajates, juhtimises, äriilgis ja muus olulises teabes kindla aja jooksul.

Edu kriteeriumid:

1. Programm, mis suudab lugeda erinevalt vormistatud aruannetest põhilisi äriandmeid.
2. Aja kokkuhoid ettevõtte ülevaate koostamisel.

Ressurssid:

1. e-Äriregister
2. OpenAI GPT

Nõuded, eeldused ja piirangud:

Nõuded:

1. Võimekus eristada ja otsida andmeid erineva struktuuriga aruannetest.

Eeldused:

1. Ligipääs aastaaruannetele.

Piirangud:

1. Tasuline ligipääs e-Äriregistri API-le.

Riskid:

1. Suured erinevused aastaaruannete vormingus või struktuuris.

2. Ebatäpne andmete klassifikatsioon aruandes.

Terminoloogia:

Aines õpitud erialane terminoloogia ja põhiline äriterminoloogia.

Kulu ja kasu:

Kulu: Programmi arendamine, testimine ja hooldamine.

Kasu: Aruannete koostamisele kulunud aja kokkuhoid.

Andmekaeve eesmärgid:

1. Tuvastada ja eraldada tähtsad finantsnäitajad ettevõtte aastaaruandest.

Andmekaeve edu kriteeriumid:

1. Võimekus tuvastada otsitavaid parameetreid erineva struktuuriga PDF-idest.
2. Täpsus otsitud andmete tuvastamisel ja võrdlemisel eelmiste aastatega.

ANDMED

Andmete kogumine

1. Struktureeritud andmed

Tavaliselt on aruannetes tabelid kus on välja toodud kõige olulisemad näitajad.

2. Mittestruktureeritud andmed

Enamus aruannet on kirjutatud vabas vormis tekstiväljadena. Seal leidub ettevõttele spetsiifiline info.

Andmete olemasolu

Andmed on olemas e-Äriregistris, kust saab aastaaruandeid tasuta allalaadida. Olemas on ka API, kuid selle kasutamine on tasuline.

Andmete valik

Kõikidel ettevõtetel on samad põhilised finantsnäitajad (kasum, käive jne). Ühe ettevõtte puhul võrreldakse näidajaid mitme aasta (aruande) vältel.

Andmete kirjeldamine

Nagu enne mainitud on andmed kas struktureeritud tabelites või tekstiväljadena. Tabelid on ettevõtete lõikes sarnased ja seega lihtsamini võrreldavad. Tabelites kajastuvad ettevõtte finantsnäitajad arvuliselt. Tavaliselt kasutatakse ühikuks tuhat eurot.

PLANEERIMINE

Põhilised etapid on aastaaruannete automaatne kogumine äriregistrist, aruannetest vajaliku informatsiooni eraldamine, eraldatud andmete analüüs ja ettevõtte lõpliku aruande koostamine. Kõige keerulisem on aruannetest vajalike andmete kättesaamine. Andmete otsimiseks ja

eraldamiseks aruannetest kasutame tõenäoliselt OpenAI GPT-d. Kui andmed on käes, siis neid analüüsitakse ning koostatakse mingi Pythoni paketiga ülevaatlik PDF. Esialgu kindlat töö ja ajajaotust liikmete vahel ei ole.