

UNIVERZA V LJUBLJANI  
EKONOMSKA FAKULTETA

MAGISTRSKO DELO

**DEJAVNIKI ESG IN FINANČNA USPEŠNOST PODJETIJ**

Ljubljana, september 2022

JAN ŠTIHEC

## IZJAVA O AVTORSTVU

Podpisani Jan Štihec, študent Ekonomske fakultete Univerze v Ljubljani, avtor predloženega dela z naslovom Dejavniki ESG in finančna uspešnost podjetij, pripravljenega v sodelovanju s svetovalcem izr. prof. dr. Igorjem Lončarskim

### IZJAVLJAM

1. da sem predloženo delo pripravil samostojno;
2. da je tiskana oblika predloženega dela istovetna njegovi elektronski obliki;
3. da je besedilo predloženega dela jezikovno korektno in tehnično pripravljeno v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani, kar pomeni, da sem poskrbel, da so dela in mnenja drugih avtorjev oziroma avtoric, ki jih uporabljam oziroma navajam v besedilu, citirana oziroma povzeta v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani;
4. da se zavedam, da je plagiatorstvo – predstavljanje tujih del (v pisni ali grafični obliki) kot mojih lastnih – kaznivo po Kazenskem zakoniku Republike Slovenije;
5. da se zavedam posledic, ki bi jih na osnovi predloženega dela dokazano plagiatorstvo lahko predstavljalo za moj status na Ekonomski fakulteti Univerze v Ljubljani v skladu z relevantnim pravilnikom;
6. da sem pridobil vsa potrebna dovoljenja za uporabo podatkov in avtorskih del v predloženem delu in jih v njem jasno označil;
7. da sem pri pripravi predloženega dela ravnal v skladu z etičnimi načeli in, kjer je to potrebno, za raziskavo pridobil soglasje etične komisije;
8. da soglašam, da se elektronska oblika predloženega dela uporabi za preverjanje podobnosti vsebine z drugimi deli s programsko opremo za preverjanje podobnosti vsebine, ki je povezana s študijskim informacijskim sistemom članice;
9. da na Univerzo v Ljubljani neodplačno, neizključno, prostorsko in časovno neomejeno prenašam pravico shranitve predloženega dela v elektronski obliki, pravico reproduciranja ter pravico dajanja predloženega dela na voljo javnosti na svetovnem spletu preko Repozitorija Univerze v Ljubljani;
10. da hkrati z objavo predloženega dela dovoljujem objavo svojih osebnih podatkov, ki so navedeni v njem in v tej izjavi.

V Ljubljani, dne \_\_\_\_\_

Podpis študenta: \_\_\_\_\_

# KAZALO

<b>UVOD</b> .....	<b>1</b>
<b>1 DEJAVNIKI ESG IN FINANČNA USPEŠNOST PODJETIJ</b> .....	<b>3</b>
<b>1.1 Kaj so dejavniki ESG</b> .....	<b>3</b>
<b>1.2 Finančna uspešnost podjetij</b> .....	<b>5</b>
<b>1.3 Pregled ugotovitev obstoječe literature</b> .....	<b>6</b>
<b>2 RAČUNALNIŠKE METODE ZA ANALIZO TEKSTA</b> .....	<b>7</b>
<b>2.1 Pridobivanje in priprava podatkov</b> .....	<b>8</b>
2.1.1 Razčlemba besedila v povedi .....	8
2.1.2 Priprava povedi za analizo.....	8
<b>2.2 Metode strojnega učenja pri rudarjenju besedil</b> .....	<b>10</b>
2.2.1 Nenadzorovano strojno učenje .....	10
2.2.2 Nadzorovano strojno učenje .....	10
<b>2.3 Pretvorba teksta v računalniku razumljivo obliko</b> .....	<b>11</b>
2.3.1 Term frequency – inverse document frequency .....	11
2.3.2 Word Embedding.....	12
2.3.3 ELMo in BERT .....	13
<b>2.4 Analiza sentimenta ali analiza razpoloženja</b> .....	<b>13</b>
<b>3 RAZISKAVA POVEZAVE MED DEJAVNIKI ESG IN FINANČNO USPEŠNOSTJO PODJETIJ</b> .....	<b>14</b>
<b>3.1 Opredelitev raziskovalnih vprašanj in metodologija</b> .....	<b>14</b>
<b>3.2 Izbira podjetij</b> .....	<b>15</b>
<b>3.3 Kazalniki finančne uspešnosti</b> .....	<b>17</b>
<b>3.4 ESG ocena podjetij</b> .....	<b>19</b>
<b>3.5 Priprava podatkov</b> .....	<b>20</b>
3.5.1 Prenos letnih poročil izbranih 100 podjetij .....	20
3.5.2 Ekstrakcija povedi iz poročil .....	20
3.5.3 Ekstrakcija ESG povedi.....	21
3.5.4 Analiza sentimenta ali razpoloženja .....	22
3.5.5 Izračun kazalnikov finančne uspešnosti .....	24
<b>3.6 Metodologija izračuna povprečne korelacije</b> .....	<b>24</b>
<b>4 REZULTATI</b> .....	<b>25</b>

<b>4.1</b>	<b>Primerjava največjih 50 in najmanjših 50 podjetij iz skupine S&amp;P 500.....</b>	<b>25</b>
4.1.1	Primerjava finančnih kazalnikov.....	25
4.1.2	Primerjava tržnih kazalnikov .....	29
4.1.3	Primerjava rezultatov analize teksta.....	31
<b>4.2</b>	<b>Primerjava najboljših 50 in najslabših 50 podjetij po ESG oceni .....</b>	<b>35</b>
4.2.1	Primerjava finančnih kazalnikov.....	37
4.2.2	Primerjava tržnih kazalnikov .....	41
<b>4.3</b>	<b>Analiza korelacije med ESG oceno in finančno uspešnostjo podjetij.....</b>	<b>43</b>
<b>4.4</b>	<b>Primerjava med gospodarskimi panogami .....</b>	<b>45</b>
<b>4.5</b>	<b>Diskusija .....</b>	<b>51</b>
<b>4.6</b>	<b>Omejitve, izboljšave in priporočila za nadaljnje raziskave.....</b>	<b>53</b>
	<b>SKLEP.....</b>	<b>54</b>
	<b>LITERATURA IN VIRI.....</b>	<b>56</b>

## KAZALO TABEL

Tabela 1: ESG kvalitete po dejavnikih ESG .....	4
Tabela 2: Skupini največjih 50 in najmanjših 50 podjetij glede na povprečno tržno kapitalizacijo od leta 2017-2021 .....	16
Tabela 3: Prikaz napovedi ESG napovednega modela.....	22
Tabela 4: Prikaz napovedi modela BERT za klasifikacijo razpoloženja .....	23
Tabela 5: Povprečna korelacija med razpoloženjem v povedih, ki se navezujejo na ESG in tistih, ki se ne.....	35
Tabela 6: Skupini najboljših 50 in najslabših 50 podjetij glede na povprečno ESG oceno	36
Tabela 7: Matrika povprečnih korelacijskih koeficientov med ESG oceno, posameznimi deli ESG ocene in kazalniki finančne uspešnosti.....	43
Tabela 8: GICS panoge .....	45
Tabela 9: Število pripadajočih podjetij za vsako GICS panogo.....	45
Tabela 10: Točke vsake panoge glede na mero ESG ocene.....	47
Tabela 11: Panoge razvrščene po oceni ESG sentimenta .....	48

## KAZALO SLIK

Slika 1: Ilustracija delovanja modelov CBOW in Skip-Gram .....	12
Slika 2: HTML elementi in pripadajoči slogovni atributi .....	21
Slika 3: Primerjava rasti prihodkov največjih 50 in najmanjših 50 podjetij .....	25

Slika 4: Primerjava bruto marže največjih 50 in najmanjših 50 podjetij.....	26
Slika 5: Primerjava neto marže največjih 50 in najmanjših 50 podjetij .....	26
Slika 6: Primerjava marže prostega denarnega toka največjih 50 in najmanjših 50 podjetij .....	27
Slika 7: Primerjava marže iz poslovanja največjih 50 in najmanjših 50 podjetij.....	27
Slika 8: Primerjava ROA največjih 50 in najmanjših 50 podjetij .....	28
Slika 9: Primerjava ROE največjih 50 in najmanjših 50 podjetij.....	28
Slika 10: Primerjava povprečne skupne vrednosti finančnih kazalnikov največjih 50 in najmanjših 50 podjetij .....	29
Slika 11: Primerjava povprečne letne donosnosti delnic največjih 50 in najmanjših 50 podjetij.....	30
Slika 12: Primerjava povprečnega standardnega odklona donosnosti delnic največjih 50 in najmanjših 50 podjetij .....	30
Slika 13: Primerjava variance donosnosti delnic največjih 50 in najmanjših 50 podjetij ...	31
Slika 14: Primerjava povprečne ESG ocene največjih 50 in najmanjših 50 podjetij .....	31
Slika 15: Primerjava povprečnega števila ESG povedi v poročilu največjih 50 in najmanjših 50 podjetij.....	32
Slika 16: Primerjava povprečnega odstotka ESG povedi v poročilu največjih 50 in najmanjših 50 podjetij .....	32
Slika 17: Primerjava povprečnega sentimenta celotnega poročila največjih 50 in najmanjših 50 podjetij.....	33
Slika 18: Primerjava povprečnega sentimenta celotnega poročila brez nevtralnih povedi največjih 50 in najmanjših 50 podjetij.....	33
Slika 19: Primerjava povprečnega sentimenta ESG povedi največjih 50 in najmanjših 50 podjetij.....	34
Slika 20: Primerjava povprečnega sentimenta ESG povedi brez nevtralnih povedi največjih 50 in najmanjših 50 podjetij .....	34
Slika 21: Primerjava rasti prihodkov najboljših 50 in najslabših 50 podjetij glede na ESG oceno.....	38
Slika 22: Primerjava bruto marže najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	38
Slika 23: Primerjava neto marže najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	39
Slika 24: Primerjava marže prostega denarnega toka najboljših 50 in najslabših 50 podjetij glede na ESG oceno.....	39
Slika 25: Primerjava marže iz poslovanja najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	40
Slika 26: Primerjava ROA najboljših 50 in najslabših 50 podjetij glede na ESG oceno ....	40
Slika 27: Primerjava ROE najboljših 50 in najslabših 50 podjetij glede na ESG oceno.....	41
Slika 28: Primerjava povprečne skupne vrednosti finančnih kazalnikov najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	41

Slika 29: Primerjava povprečne letne donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	42
Slika 30: Primerjava povprečnega odklona donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	42
Slika 31: Primerjava variance donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno .....	43
Slika 32: Razvrstitev gospodarskih dejavnosti po E in S tveganju .....	46
Slika 33: Skupna povprečna ESG ocena za vsako GICS panogo .....	47
Slika 34: ESG ocena za vsako panogo po analiziranih letih .....	50

## SEZNAM KRATIC

angl. – angleško

**API** – (angl. Application Programming Interface); aplikacijski programski vmesnik

**CFP** – (angl. Corporate Financial Performance); finančna uspešnost podjetij

**CSP** – (angl. Corporate Social Performance); uspešnost podjetij na področju družbene odgovornosti

**CSR** – (angl. Corporate Social Responsibility); družbena odgovornost podjetij

**EBIT** – (angl. Earnings Before Interest and Taxes); dobiček pred obrestmi in davki

**IT** – Informacijska Tehnologija

**NLP** – (angl. Natural Language Processing); obdelava naravnega jezika

**PMS** – (angl. Performance Measurement Systems); sistemi za merjenje uspešnosti

**ROA** – (angl. Return On Assets); donosnost sredstev

**ROE** – (angl. Return On Equity); donosnost lastniškega kapitala

**ROIC** – (angl. Return On Invested Capital); donosnost vloženega kapitala

**ROS** – (angl. Return On Sales); donosnost prodaje

**SRI** – (angl. Socially Responsible Investing); družbeno odgovorno investiranje

**SVM** – (angl. Support-Vector Machine); metoda podpornih vektorjev

**S&P 500** – borzni indeks Standard and Poor's 500

**TF-IDF** – (angl. Term Frequency – Inverse Document Frequency)

## UVOD

Trajnostni razvoj in družbena odgovornost sta besedni zvezi, ki ju v zadnjih letih oziroma celo desetletjih slišimo ali preberemo vse bolj pogosto. Izhajata iz zavedanja, da je potrebno v želji po čim hitrejšem napredku in razvoju misliti tudi na negativne posledice, ki jih morda ne občutimo v realnem času, bodo pa sigurno vplivale na naš način življenja v prihodnosti.

Kaj pravzaprav sta trajnostni razvoj in družbena odgovornost. Svetovna komisija za okolje in razvoj postavi naslednjo definicijo (Brundtland, 1987):

*»Trajnostni razvoj je razvoj, ki zadovoljuje sedanje potrebe, brez da bi ogrožal zmožnost prihodnjih generacij, da zadovoljijo svoje potrebe.«*

Široka definicija, ki lahko služi kot smernica za odločanje in upravljanje v različnih družbenih sferah. V magistrskem delu se omejujem na ekonomsko sfero. Še širši pojem kot trajnostni razvoj je družbena odgovornost. Ta govori o načelu, da moramo posamezniki in podjetja delovati v dobro okolja in družbe kot celote. Evropska komisija definira družbeno odgovornost podjetij (angl. corporate social responsibility – CSR) kot (European Commission, 2001):

*»Proces, v katerem podjetja v svoje poslovne dejavnosti in odnose z interesnimi skupinami vključujejo družbena, okoljska in etična vprašanja ter vprašanja človekovih pravic.«*

V ekonomski sferi se za različne vidike in vprašanja znotraj zgoraj opisanih načel pogosto uporablja kratica **ESG**. Ta je sestavljena iz treh dejavnikov (The Global Compact, 2004):

- **Environmental** – okoljski dejavniki.
- **Social** – družbeni dejavniki.
- **Governance** – upravljavski dejavniki.

Če je z vidika družbe kot celote vprašanje trajnostnega razvoja zelo široko, je z vidika investitorjev pogosto bolj usmerjeno: *ali obstaja povezava med zavezanostjo k trajnosti in finančno uspešnostjo?*

Da bi lahko odgovorili na to vprašanje, je najprej potrebno določiti, kako izmeriti ali oceniti trajnostno naravnost in družbeno odgovornost podjetja. Dejavniki ESG predstavljajo enega od možnih načinov. Vprašanje postane, kako izmeriti, kako dobro podjetje izpolnjuje ESG kriterije in sledi načelom družbeno odgovornega delovanja.

Enega od odgovorov na to vprašanje ponuja informacijska tehnologija (v nadaljevanju IT). Ta nam dandanes omogoča, da kvantificiramo sicer kvalitativne podatke. Daje nam možnost, da iz za človeka prevelike količine podatkov pridobimo informacije in znanje, ki nam lahko pomaga k boljšem razumevanju določene problematike in pri sprejemanju odločitev.

Magistrsko delo predstavlja praktično uporabo znanja pridobljenega tekom dodiplomskega študija na Fakulteti za računalništvo in informatiko ter magistrskega študija poslovna informatika na Ekonomski fakulteti. Magistrsko delo prav tako nadaljuje raziskovalno delo, ki sem ga tekom študija opravljal v sodelovanju z Inštitutom Jožef Stefan pri projektih Vpliv formalnih in neformalnih komunikacij na trg kapitala in Kvantitativna in kvalitativna analiza nereguliranih delov finančnega poročanja podjetij.

Delo združuje področja računalništva, ekonomije in lingvistike ter aplicira računalniške rešitve na reševanje definirane družbene problematike. Pod-področje računalništva, ki vključuje rešitve predstavljene in uporabljene v tem delu, je podatkovna znanost ali tudi znanost o podatkih. To je področje, ki se, zahvaljujoč vedno večjim zmogljivostim računalnikom, širi izjemno hitro. Če sem na ekonomski strani tega dela kot pogosto slišani oziroma prebrani besedni zvezi navedel trajnostni razvoj in družbena odgovornost, lahko na računalniški strani navedem pojma strojno učenje in umetna inteligenca. Napredki strojnega učenja nam omogočajo, da podatke analiziramo in razumemo s pomočjo pametnih računalnikov ali sistemov. S hitrim razvojem se večja tudi nabor orodij in tehnologij, ki jih lahko uporabimo za reševanje našega problema. Velik izziv tako postane ustrezna izbira in uporaba sicer zelo močnih orodij in rešitev.

Namen magistrskega dela je z interdisciplinarnim pristopom, ki združuje ekonomska znanja in uporabo naprednih računalniških tehnologij, doprinesiti k boljšemu razumevanju povezave med finančno uspešnostjo poslovanja podjetij in njihovo usmerjenostjo v trajnostno poslovanje (merjeno z ESG oceno).

Da bom ta namen lahko dosegel, definiram sledeče cilje magistrskega dela:

- Raziskati in predstaviti ugotovitve sorodnih študij.
- Predstaviti uporabo računalniških tehnologij za analizo teksta in pridobivanje informacij.
- Razviti algoritem za pridobivanje in obdelavo podatkov s spleta.
- Razviti algoritem za pridobivanje informacij iz delno strukturiranih besedil.
- Razviti in evalvirati algoritem za ocenjevanje razpoložjenja v povedih.
- Definirati in na podlagi analize teksta izračunati ESG oceno podjetij.
- Z uporabo statističnih metod raziskati povezavo med ESG oceno in finančno uspešnostjo podjetij.
- Raziskati povezavo med izračunano ESG oceno in ESG tveganjem.

V prvem poglavju se osredotočim na del, ki je povezan s finančno analizo oz. merjenjem finančne uspešnosti. Pri pregledu strokovne literature celostno zajamem ugotovitve študij zadnjih 40 let, ki so raziskovale povezavo med družbeno odgovornim delovanjem in finančno uspešnostjo podjetij. Ugotovitve dosedanjih študij pripomorejo k boljšemu razumevanju problematike, hkrati pa jih uporabim za opredelitev raziskovalnih vprašanj in primerjavo z ugotovitvami moje raziskave. Dodatno poskušam pri pregledu literature



ugotoviti, s katerimi vidiki finančne uspešnosti je še posebej povezano družbeno odgovorno delovanje in dejavniki ESG.

Drugo poglavje se posveti podatkovni znanosti, kjer predstavim obstoječe metode in tehnologije podatkovne analize. Metode so predstavljene v enakem zaporedju, kakor so si sledile tudi pri samem razvoju za analizo potrebnih algoritmov. Cilj poglavja je ne samo opisati izbrane metode, ampak predstaviti problematiko širše ter argumentirati izbiro uporabljenih orodij. Poglavje bralcu omogoča uporabo enakega ali podobnega postopka za različne sorodne analize.

Tretje poglavje vsebuje empirični del. Najprej opredelim raziskovalna vprašanja in predstavim metodologijo analize ter izbiro podjetij. Sledi definicija kazalnikov finančne uspešnosti in ESG ocene, kar v nadaljevanju uporabim za raziskovanje povezave med dejavniki ESG in finančno uspešnostjo. Nato predstavim razvoj algoritmov in način pridobivanja ter obdelave vseh potrebnih podatkov.

Četrto poglavje najprej grafično predstavi rezultate analize in primerjavo skupin podjetij. V tem poglavju s pomočjo statističnih metod raziskujem definirana raziskovalna vprašanja. Poglavje zaključim z diskusijo o pridobljenih rezultatih in ugotovitvah.

V sklepnem poglavju ovrednotim dosego zastavljenih ciljev ter predstavim doprinos magistrskega dela raziskovani problematiki.

## **1 DEJAVNIKI ESG IN FINANČNA USPEŠNOST PODJETIJ**

V zadnjih letih se vse bolj poudarja pomen družbeno odgovornega investiranja (angl. socially responsible investing – SRI) ali tudi ESG investiranja (Sciarelli, Cosimato, Landi & Iandolo, 2021). To daje pri odločitvah o novih investicijah velik pomen trajnostni usmerjenosti. Podjetja kot tudi investitorje zanima, ali investiranje v skladu z ESG merili prinaša tudi strateško prednost.

Da lahko raziskujemo povezavo med dejavniki ESG in finančno uspešnostjo podjetij, moramo najprej razumeti, kaj dejavniki ESG sploh so in na katere aktivnosti organizacij se navezujejo.

### **1.1 Kaj so dejavniki ESG**

Dejavniki ESG so način kako lahko izrazimo in merimo družbeno odgovornost podjetja (Sassen, Hinze & Hardeck, 2016). ESG družbeno odgovornost razdeli na okoljski vidik, družbeni vidik in upravljavski vidik. Ta delitev nam omogoča, da bolj podrobno analiziramo delovanje podjetja z ozirom na trajnostni razvoj in delovanje, ki družbi koristi.

**Okoljski dejavniki (E)** vključujejo vplive na okolje, ki so rezultat aktivnosti podjetja. V to kategorijo spadajo vprašanja o ravnanju z odpadki, onesnaževanju, podnebnih spremembah, porabi energije oziroma varčnosti, izpustih, ohranjanju narave in varovanju živali (Investopedia, 2022).

**Družbeni dejavniki (S)** se nanašajo na vprašanja o odnosu podjetja z deležniki (Fitzpatrick, Neilan & Reilly, 2020). Pod ta dejavnik vključujemo skrb za varnost pri delu, splošna skrb za zdravje in počutje zaposlenih, varnost končnih produktov za kupce, etičen odnos do kupcev ali strank, zahtevanje izpolnjevanja standardov od dobaviteljev, donacije, prostovoljstva (Investopedia, 2022)

**Upravljavski dejavniki (G)** zajemajo vprašanja o procesu odločanja v organizaciji, delitvi pravic, pooblastil in odgovornosti ter namen in vloga upravljavskih funkcij (S&P Global, 2020)

Za še boljše razumevanje dejavnikov ESG in izpolnjevanja kriterijev ESG v tabeli 1 navedem sezname nekaterih ESG kvalitete, ki jih predstavi podjetje Trillium Asset Management (2022), ki se ukvarja z družbeno odgovornim ali ESG investiranjem.

*Tabela 1: ESG kvalitete po dejavnikih ESG*

<b>Okolje</b>	<b>Družba</b>	<b>Upravljanje</b>
<ul style="list-style-type: none"> <li>• Omejevanje škodljivih onesnažilcev in kemikalij</li> <li>• Odgovorno ravnanje s surovinami in materiali</li> <li>• Aktivno zmanjševanje emisij toplogrednih plinov</li> <li>• Uporaba obnovljivih virov energije</li> </ul>	<ul style="list-style-type: none"> <li>• Zaposleni prejemajo pošteno plačo</li> <li>• Zanesljiva strategija upravljanja človeškega kapitala</li> <li>• Spodbujanje raznolikosti in vključenosti</li> <li>• Podpiranje pravic LGBTQ</li> <li>• Etična dobavna veriga</li> </ul>	<ul style="list-style-type: none"> <li>• Prizadevanje za raznolikost v upravnih odborih</li> <li>• Ločitev funkcij izvršni direktor in predsednik upravnega odbora</li> <li>• Razumne plače vodstvenih delavcev</li> <li>• Predanost korporativni transparentnosti</li> </ul>

*Prirejeno po Trillium Asset Management (2022).*

Hitro lahko opazimo, da je zelo težko definirati in izmeriti kako uspešno je določeno podjetje na področju ESG. Gospodarske panoge so izpostavljenih drugačnim ESG tveganjem glede na dejavnosti s katerimi se pripadajoča podjetja ukvarjajo (Pástor, Stambaugh & Taylor, 2021; Xiong, 2021). Prav tako se med panogami razlikuje dojemanje dejavnikov ESG in relevantnost za poslovne dejavnosti (Sassen, Hinze & Hardeck, 2016). Nadalje Eccles,

Krzus, Rogers in Serafeim (2012) trdijo, da mora biti ocenjevanje in poročanje o dejavnikih ESG razvito za vsako panogo posebej. Vse bolj pogosto se izraža potreba po vzpostavitvi okvira za sporočanje ESG prizadevanj podjetij, ki bi investitorjem dal vse relevantne informacije (Fornasari, 2020). Fornasari (2020) nadaljuje, da vprašanja, kaj meriti in kako meriti, ni samo tehničen izziv, temveč je odgovor odvisen tudi od političnih odločitev, kar je v skladu z ugotovitvami Sassen, Hinze in Hardeck (2016). Fornasari (2020) še doda, da so si koncepti in dojetje trajnosti različni, kar vpliva na izbor vključenih indikatorjev. Z ocenjevanjem ESG delovanja se ukvarjajo različne institucije, ki podajo svoje ESG ocene podjetij (recimo Bloomberg in Fitch Ratings). Običajno na to oceno najbolj vpliva uspešnost podjetja pri obvladovanju izzivov na področju ESG.

Z vse večjim poudarkom na družbeno odgovornost in trajnostni razvoj so ESG ocene postale za podjetja pomemben faktor, ki vpliva na strateške odločitve (Dofleitner, Halbritter & Nguyen, 2015). Vse večjo željo po družbeno odgovornem delovanju izražajo tudi deležniki, ki od managerjev pričakujejo vpeljavo rešitev, ki so v skladu z ESG smernicami (Crilly, Zollo & Hansen, 2012; Egorova, Grishunin & Karminsky, 2022). Prav tako višje zavedanje o pomenu trajnosti med regulatorji, potrošniki in investitorji prinaša višje ESG tveganje s katerim se podjetja soočajo (Hübel & Scholz, 2020). Način kako izmeriti ESG ravnanje podjetij je tako zelo razširjena problematika, s katero se ukvarjajo tako znanost kot tudi podjetja sama.

## **1.2 Finančna uspešnost podjetij**

Eden od glavnih ciljev podjetij je jasno izboljšati finančno uspešnost. Finančno uspešnost merimo s pomočjo določenih finančnih kazalnikov. Tudi tukaj pa se poraja vprašanje, kateri so ključni kazalniki, ki nam povedo največ o finančni uspešnosti. Dossi in Patelli (2010) predstavita seznam najbolj pogosto poročenih finančnih kazalnikov v sistemih za merjenje uspešnosti (angl. performance measurement systems – PMS) v multinacionalnih podjetjih. Pri vrhu najbolj pogostih uporabljenih finančnih kazalnikov so: prihodki od prodaje, prihodki iz poslovanja, prispevna marža, bruto marža, neto prihodki in denarni tok. Na seznamu najbolj pogosto poročenih meril finančne uspešnosti iz raziskave avtorjev Boaventura, Silva in Bandeira-de-Mello (2012) so: donosnost sredstev (angl. return on assets, v nadaljevanju ROA), donosnost lastniškega kapitala (angl. return on equity, v nadaljevanju ROE), rast prodaje, donosnost prodaje (angl. return on sales, v nadaljevanju ROS) in prispevna marža. K finančni uspešnosti podjetij lahko vključimo tudi kazalnike izračunane na podlagi tržnih podatkov, kot so donosnost delnic in volatiliteta izražena z odklonom in varianco donosnosti. Temu delu zelo sorodna študija avtorjev Nollet, Filis in Mitrokostas (2016), ki analizira povezavo med ESG in različnimi kazalniki finančne uspešnosti pri podjetjih iz skupine S&P 500 med leti 2007 in 2011, vključuje kazalnike ROA, donosnost vložene kapitala (angl. return on invested capital - ROIC) in pričakovano donosnost delnic. V analizo vključene kazalnike finančne uspešnosti podjetij predstavim v poglavju 3.3.

### 1.3 Pregled ugotovitev obstoječe literature

V zadnjih desetletjih je raziskav povezave med uspešnostjo na področju družbene odgovornosti (angl. corporate social performance, v nadaljevanju CSP) in finančno uspešnostjo podjetij (angl. corporate financial performance, v nadaljevanju CFP) vedno več (Friede, Busch & Bassen 2015). V 90ih nastane Porterjeva hipoteza imenovana po ekonomistu Michaelu Porterju, ki pravi, da lahko višji okoljski standardi in regulacije spodbudijo podjetja k hitrejši inovaciji, boljšemu in učinkovitejšemu upravljanju z viri in s tem pripomorejo k višji konkurenčnosti (Porter & Linde, 1995). Ta hipoteza je spodbudila ogromno debat in raziskav, ki so na svoj način poskušale hipotezo potrditi ali ovreči. Ugotovitve teh raziskav niso povsem enoznančne. K zahtevnosti problematike doprinese dejstvo, da je družbena odgovornost nefinančni dejavnik, ki nima določenega načina merjenja, o čemer pišem že v poglavju 1.1. Dejavniki ESG predstavijo tri različne vidike, vendar pa ne določijo točnega načina merjenja uspešnosti podjetij pri obvladovanju teh vidikov. ESG ocene, ki jih objavljajo institucije, so pridobljene skozi različne analize in postopke. Prav tako je zelo težko govoriti o neposrednem vplivu uspešnega družbeno odgovornega delovanja na finančno uspešnost. Klub vsemu naštetemu večina študij zadnjih štirih desetletij kaže, da obstaja pozitivna ESG-CFP povezava (Friede, Busch & Bassen, 2015; Garcia & Orsato, 2020). Orlitzky, Schmidt in Rynes (2003) skozi metaanalizo 52 študij ugotovijo, da obstaja pozitivna korelacija med CSP in CFP in trdijo, da je povezava dvosmerna. Po še večji metaanalizi 167 raziskav Margolis, Elfenbein in Walsh (2009) zaključijo, da obstaja pozitivna a majhna povezava med CSP in CFP in da je moč povezave odvisna od specifične dimenzije CSP. Van Beurden in Gössling (2008) na podlagi metaanalize ugotovita, da obstaja jasna pozitivna korelacija med CSP in CFP ter da so študije, ki kažejo nasprotno, zastarele, saj se je družba spremenila. Tezo o pozitivni povezavi potrjuje tudi Shinohara (2014), ki zaključi, da pozitivna povezava obstaja in da je še močnejša po petih letih – torej da se CSP uspešnost najmočnejše odraža v CFP uspešnosti čez pet let. Barnett in Salomon (2012) predstavita CSP-CFP povezavo kot bolj kompleksno in sicer, da vlaganje v uvajanje družbeno odgovornih praks lahko privede do boljših odnosov z deležniki in s tem upraviči svojo ceno, a hkrati prinese obvezo, da podjetje te boljše odnose pretvori v strateško prednost. V nasprotnem primeru je lahko CSP-CFP povezava negativna. Tej ugotovitvi se pridružujejo Nollet, Filis in Mitrokostas (2016), ki zaključijo, da je povezava CSP in CFP pozitivna šele, ko je presežena določena točka vlaganja v CSP. Nadaljujejo, da so za dosego pozitivne povezave potrebna dolgoročna vlaganja v CSP in znatna sredstva ter da to velja samo za upravljavski vidik ESG. Friede, Busch in Bassen (2015) analizirajo več kot 2000 empiričnih študij od leta 1970 do leta 2015 in zaključijo, da 90% raziskav najde nenegativno ESG-CFP povezavo. Tri leta pozneje Busch in Friede (2018) na podlagi meta analize drugega reda zaključita, da je CSP-CFP povezava pozitivna, dvosmerna, robustna in značilna. Busch in Friede (2018) nadaljujeta, da je povezava pozitivna ne glede na to, ali se podjetje osredotoča na ekološke ali družbene vidike. Sassen, Hinze in Hardeck (2016) poročajo o vplivu višje ESG ocene na nižje skupno in specifična tveganja. Ashwin Kumar in drugi (2016) ugotovijo, da obstaja povezava med dejavniki ESG

in vrednostmi delnic - bolj natančno povezava z nižjo volatilitnostjo delnic. Poznejša študija avtorjev La Torre, Mango, Cafaro in Leo (2020), ki se osredotoči na Eurostoxx50 podjetja, pride do zaključka, da zavezanost sledenju ESG smernicam ne vpliva na finančno uspešnost. Dodatno ugotovitev o pozitivni povezavi izpodbijajo nekatere študije, ki se osredotočijo na podjetja iz držav v razvoju (Garcia & Orsato, 2020). Kot možen razlog za to Garcia in Orsato (2020) navedeta dejstvo, da v državah v razvoju težko sledijo visokim standardom, ki jih postavljajo svetovne organizacije. Zahtevani vložki v družbeno odgovorno delovanje lahko tako postanejo previsoki. Ker ESG v ospredje postavlja vpliv na družbo, se moramo zavedati, da so si svetovne kulture zelo različne. Prakse, ki jih na enem delu sveta sprejemajo kot pozitivne, morda niso enako sprejete na drugem delu sveta.

V tem magistrskem delu se omejim na podjetja iz ZDA, tako lahko po pregledu širše literature zaključim, da velika večina študij potrjuje pozitivno povezavo med ESG in CFP.

## **2 RAČUNALNIŠKE METODE ZA ANALIZO TEKSTA**

Računalniška analiza teksta ali besedil spada v širše področje obdelave naravnega jezika (angl. natural language processing, v nadaljevanju NLP). NLP lahko definiramo kot spekter računalniških metod za avtomatsko analizo in reprezentacijo ali generiranje človeškega jezika (Cambria & White, 2014). V magistrskem delu se omejim na analizo zapisanega človeškega jezika. To pogosto imenujemo tudi rudarjenje besedil (angl. text mining). Tekst ali besedilo v osnovni obliki računalniku ni razumljivo v smislu pomena in vsebine. NLP se ukvarja s tehnikami, ki nam omogočajo, da človeški jezik v zapisani obliki računalniku predstavimo na način, da bo lahko presojal tudi o vsebini in pomenu. Te tehnike postanejo še toliko bolj zanimive zaradi izjemne računske zmogljivosti novodobnih računalnikov. Ta računska zmogljivost nam daje možnost obdelave velike količine podatkov, kar bi sicer zahtevalo nešteto človeških ur. Hiter razvoj področja NLP je posledica zavedanja, da lahko s pomočjo tehnik NLP iz večjega števila podatkov pridobimo informacije, ki lahko povečajo naše znanje in nam pomagajo pri sprejemanju odločitev. Razvoj NLP področja stremi k cilju, da lahko računalnik naučimo razumevanja besedila, tako kot ga razume človek. Osnovni pristop k tej problematiki poskuša sestaviti algoritem, ki bo besedilo bral na način, kot ga beremo ljudje. Glavna težava tega postopka je, da ljudje ne razumemo besed samo kot zaporedje določenih črk, ampak nam beseda predstavlja dosti več. Ko preberemo besedo, jo avtomatsko povežemo s semantično podobnimi koncepti, izkušnjami in čustvi (Cambria & White, 2014). Spisati algoritem, ki bi uspešno imitiral ta proces razumevanja, je zahteven problem, ki še ni rešen. Imamo pa na voljo različne tehnike, s katerimi lahko dokaj natančno rešimo določena vprašanja razumevanja človeškega jezika. V naslednjih štirih podpoglavjih predstavim skupine teh tehnik v takšnem zaporedju, kot jih običajno uporabimo pri reševanju NLP problemov. V magistrskem delu sem analiziral podatke zapisane v angleškem jeziku, tako da je to poglavje napisano, kot da imamo opravka z angleščino. Najnaprednejše NLP tehnike in algoritmi so običajno najprej razviti za angleški jezik. Posledično je potreba po prilagoditvi na druge svetovne jezike velika.

## 2.1 Pridobivanje in priprava podatkov

Pridobivanje in priprava podatkov za analizo je prvi korak kateregakoli NLP projekta. Čeprav se to opravilo morda na prvi pogled zdi manj pomembno in nezanimivo, ponavadi zavzame več kot 65% časa celotnega projekta (Dodds, 2020). Seveda se ta čas lahko močno razlikuje od projekta do projekta, saj so si oblike in strukture osnovnih podatkov lahko zelo različne.

### 2.1.1 Razčlemba besedila v povedi

Pri tovrstnih projektih nas običajno zanimata pomen in vsebina, zato se osredotočimo na najmanjšo semantično enoto – poved (Malo, Sinha, Korhonen, Wallenius & Takala, 2014). Če je za človeka razčlemba besedila v povedi enostavna naloga, pri računalniku ni povsem tako. Poved namreč razumemo kot pomensko enoto, ki se običajno začne z veliko začetnico in konča s piko (Leemeta, 2021). Računalnik ne zna presojati, kje se konča pomenska enota, tako da se lahko zanašamo samo na pravila o začetku in koncu povedi. Vendar pa tudi med temi pravili najdemo izjeme. Poved se najpogosteje konča s piko ali katerim drugim ločilom, a ločila se lahko pojavijo tudi znotraj povedi (recimo okrajšave kot je mr.). Računalnik mora tako znati presojati ali ločilo predstavlja krajšavo ali konec povedi. Nalogo razčlemba besedila v povedi lahko zelo natančno rešimo s pomočjo Python knjižnice NLTK (Loper & Bird, 2002). Ta uporablja algoritem nenadzorovanega strojnega učenja za zaznavanje mej povedi (več o strojnem učenju napišem v poglavju 2.3). Izmerjena natančnost tega postopka je višja kot 98% (Kiss & Strunk, 2006). Ta natančnost je izmerjena na besedilih iz časopisov, ki so zapisani v računalniku prijazni obliki. Več težav z zaznavanjem povedi imamo lahko, če je tekst pridobljen iz virov, ki ne sledijo osnovnim pravilom začenjanja in zaključevanja povedi. Najpogostejši tak vir so spletne strani, kjer so povedi ločene tudi z grafičnimi elementi ali pa so zapisane z različnimi slogi. Primer računalniku manj prijaznega besedila, ko želimo pridobiti vse povedi, je recimo naslovnica magistrskega dela. To bi zgornji algoritem stežka razčlenil v povedi.

### 2.1.2 Priprava povedi za analizo

Ko je besedilo pripravljeno v obliki povedi, je naslednji korak preoblikovanje povedi z metodami predprocesiranja ali normalizacije. Cilj tega postopka je povečati zmožnost računalnika, da povezuje besede in povedi, ki so si pomensko bližje. Pri tem postopku operiramo z besedami v povedi in ne povedjo kot celoto. Razčlemba povedi na besede je manj zahteven postopek kot razčlemba besedila v povedi in ga prav tako lahko opravimo s pomočjo knjižnice NLTK. Sledijo nekatere tehnike predprocesiranja, ki sem jih uporabil tudi sam. Izbor uporabljenih tehnik predprocesiranja je seveda odvisen od definirane cilja projekta. Lahko se zgodi, da določen korak normalizacije namerno izpustimo, saj bi negativno vplival na kasnejše pridobivanje informacij.

- **Pretvorba v male črke.** Ko je besedilo že razčlenjeno na povedi, želimo celotno poved zapisati z malimi črkami. Ne želimo, da bi računalnik razločeval med pomensko enakimi besedami, ki pa so pisane z različnimi velikostmi črk. Ta postopek je še posebej pomemben, če je bilo besedilo pridobljeno s spleta, kjer so besede pogosto pisane z veliko zaradi oblikovnih ciljev avtorja.
- **Odstranitev simbolov in ločil.** Simboli in ločila ne pomagajo pri razumevanju povedi, zato jih lahko odstranimo. Seznam odstranjenih simbolov in ločil je običajno sledeč: `!"#$%&'()*+,-./:;<=>?@[\\]^_`{|}~.` V določenih primerih lahko simbole uporabimo za dodatno ločevanje besedila na manjše dele ali pridobivanje posebnih delov besedila. Tak primer so recimo ključniki (angl. hashtag), ki se začnejo z znakom # in jih lahko razumemo kot samostojne enote besedila.
- **Zapis števil z besedo ali odstranitev.** Števila običajno ne pomagajo pri razumevanju povedi, tako da jih pogosto odstranimo. V kolikor jih želimo obdržati, jih je potrebno zapisati z besedami. Če je ciljni jezik angleščina, imamo tudi za to nalogo na voljo zanesljive Python knjižnice (recimo `num2words`).
- **Odstranitev stop besed.** Stop besede (angl. stop words) lahko definiramo kot besede, ki se z enako verjetnostjo pojavijo v dokumentih (v primeru tega dela je dokument ena poved), ki za analizo niso relevantni in tistih, ki za analizo so relevantni (Willbur & Sirotkin, 1992). To so torej besede, ki nam pri nalogah NLP ne koristijo, saj ne nosijo semantične vrednosti. Primeri stop besed so recimo: a, an, the, in, of itn. Seznam stop besed v NLTK knjižnici, ki sem jo uporabil tudi sam, vključuje 40 besed. Dodatna motivacija za odstranitev teh besed je, da s tem zmanjšamo nabor vseh možnih besed v analiziranem besedilu, kar bo pripomoglo k manjši računski zahtevnosti končnega algoritma (Kannan in drugi, 2014).
- **Krnjenje in lematizacija.** Ena od zmožnosti, ki jo mora algoritem za rudarjenje besedil imeti, je združevanje besed in povedi po podobnosti. Tako lahko določimo katere informacije želimo iz besedila pridobiti, jih kar se da dobro definiramo in nato prepustimo algoritmu, da v tekstu najde ustrezne zadetke (v primeru tega dela so to povedi). Postopka, ki izboljšata ta proces pridobivanja relevantnih delov besedila, sta krnjenje in lematizacija (angl. stemming in lemmatization). S tega postopkoma dosežemo, da naš program ne bo izpustil določenih besed zaradi drugačnih končnic in hkrati zmanjšamo nabor vseh možnih besed v našem besedilu (Balakrishnan & Lloyd-Yemoh, 2014). Preprostejša in hitrejša metoda je krnjenje, ki besedi odreže končnico in pusti samo krn. Lematizacija je zahtevnejši postopek, ki poskuša najti osnovno ali slovarsko oblike besede – lemo.

V nadaljevanju je prikazan primer obdelave povedi z metodami predprocesiranja.

Poved v prvotni pridobljeni obliki:

*These measures are used internally by the Company to evaluate the performance of it's underlying operations and to assist with it's planning and forecasting of future operating results.*

Poved po uporabi prvih štirih korakov in krnjenju:

*measur use intern compani evalu perform underli oper assist plan forecast futur oper result*

Poved po uporabi prvih štirih korakov in lematizaciji:

*measure used internally company evaluate performance underlying operation assist planning forecasting future operating result*

## **2.2 Metode strojnega učenja pri rudarjenju besedil**

Morda najhitreje rastoče področje računalništva in tehnologije nasploh je strojno učenje. Ena najzgodnejših definicij strojnega učenja postavi Arthur Samuel (1959): »strojno učenje je področje preučevanja, ki daje računalnikom zmožnost učenja brez eksplicitnega programiranja«. Nekoliko pozneje postavi Mitchell (1997) sledečo definicijo: »za računalniški program pravimo, da se uči na izkušnji I glede naloge N ob meri uspešnosti U, če se uspešnost programa na nalogi N merjena z U veča z izkušnjo I«. Glavno vprašanje strojnega učenja je, kako računalnik naučiti, da se bo izboljševal sam skozi izkušnje (Jordan & Mitchell, 2015). Ena prvih zelo razširjenih rešitev iz 90ih, ki je temeljila na strojnem učenju in vplivala na praktično vsakogar, je filtracija neželene ali s tujko spam pošte (Géron, 2019). Rešitev lepo ponazori delovanje in namen strojnega učenja. Program smo naučili, da sam na podlagi besedila prepozna ali gre za verodostojno sporočilo ali le za neželjeno pošto. Strojno učenje se danes uporablja že praktično na vseh področjih, kjer poskušamo s pomočjo računalnikov optimizirati obstoječe procese. Enako velja tudi za področji NLP in rudarjenje besedil. V osnovi lahko strojno učenje razdelimo na nadzorovano in nenadzorovano.

### **2.2.1 Nenadzorovano strojno učenje**

Nenadzorovano učenje je način učenja, kjer imamo na voljo neoznačene podatke. Pri uporabi tehnik nenadzorovanega učenja se računalnik odloča brez človeškega usmerjanja. To nam omogoča različne načine analize podatkov, kot sta recimo grupiranje na podlagi določene vrednosti ali detekcija anomalij. Pri rudarjenju besedil lahko nenadzorovano učenje uporabimo za grupiranje semantično podobnih povedi.

### **2.2.2 Nadzorovano strojno učenje**

Pri nadzorovanem učenju imamo na voljo označene podatke oziroma poznamo pravilno rešitev problema. Označke ali rešitve služijo kot navodila, kako naj se naš model odloča. Model se nauči povezovanja danih vhodnih vrednosti z določeno izhodno vrednostjo ali oznako. Poznamo različne algoritme, ki jih lahko uporabimo pri izdelavi odločitvenega modela z nadzorovanim učenjem. Najbolj pogosti so naivni Bayesov klasifikator, linearna



regresija, odločitvena drevesa, nevronske mreže, metoda podpornih vektorjev itn. (Fumo, 2017). Slednjega sem uporabil tudi sam v svoji raziskavi.

## 2.3 Pretvorba teksta v računalniku razumljivo obliko

Po tem ko je vhodno besedilo pripravljeno za nadaljnjo analizo, je potrebno določiti način predstavitve teksta v računalniku razumljivi obliki. Ta postopek se imenuje vektorizacija in z njim tekst pretvorimo v zapis s števili oziroma vektorji (Singh & Shashi, 2019). Pri večini projektov rudarjenja besedila je ta postopek nujen. Najbolj preprost način za predstavitev besedila numerično je, da vsako besedo v našem besedilu zapišemo z določenim številom. Ta metoda je seveda neučinkovita in nam ne pove nič o pomembnosti ali relevantnosti posameznih besed. Obstajajo boljši algoritmi z vidika učinkovitosti in rezultatov. Spodaj je predstavljen nekaj pristopov, ki jih lahko uporabimo za vektorizacijo teksta. Izbira je odvisna od ciljev nadaljnega dela s podatki.

### 2.3.1 Term frequency – inverse document frequency

V programih za pridobivanje informacij oziroma za rudarjenje besedil je algoritem Term frequency – inverse document frequency (v nadaljevanju TF-IDF) eden od najpogosteje uporabljenih (Aizawa, 2003). Uporabim ga tudi sam v tej raziskavi. Algoritem je možno prilagoditi glede na obliko vhodnih podatkov in cilj analize. TF-IDF vsaki besedi ali izrazu določi utež glede na pomembnost besede za pridobivanje informacij. Ta utež je določena na podlagi frekvence pojavljanja v našem besedilu. Besede, ki se pojavljajo zelo pogosto, nosijo manjšo pomembnost, saj nam o povedi ne povedo kaj dosti. Ena od pogostih nalog v rudarjenju besedil je združevanje povedi glede na pomen. Besede, ki se pojavljajo v večini povedi, nam pri tem ne pomagajo. Za primer lahko uporabim besedo »the«, ki se pojavi v skorajda vsaki povedi in o vsebini ne pove praktično nič (to besedo celo povsem odstranimo, če uporabimo katerega od algoritmov za odstranitev stop besed). Na drugi strani nosijo besede, ki se pojavljajo manj pogosto, visoko utež, saj povedo dosti o vsebini povedi in nam zelo koristijo pri razlikovanju. Vrednost TF-IDF lahko izračunamo z enačbo (1):

$$tfidf(b, p, B) = tf(b, p) * idf(b, B) \quad (1)$$

Kjer je  $b$  dana beseda,  $p$  dana poved in  $B$  celotno besedilo.  $tf(b, p)$  nam pove kolikokrat se beseda pojavi v povedi,  $idf$  pa lahko izračunamo z enačbo (2):

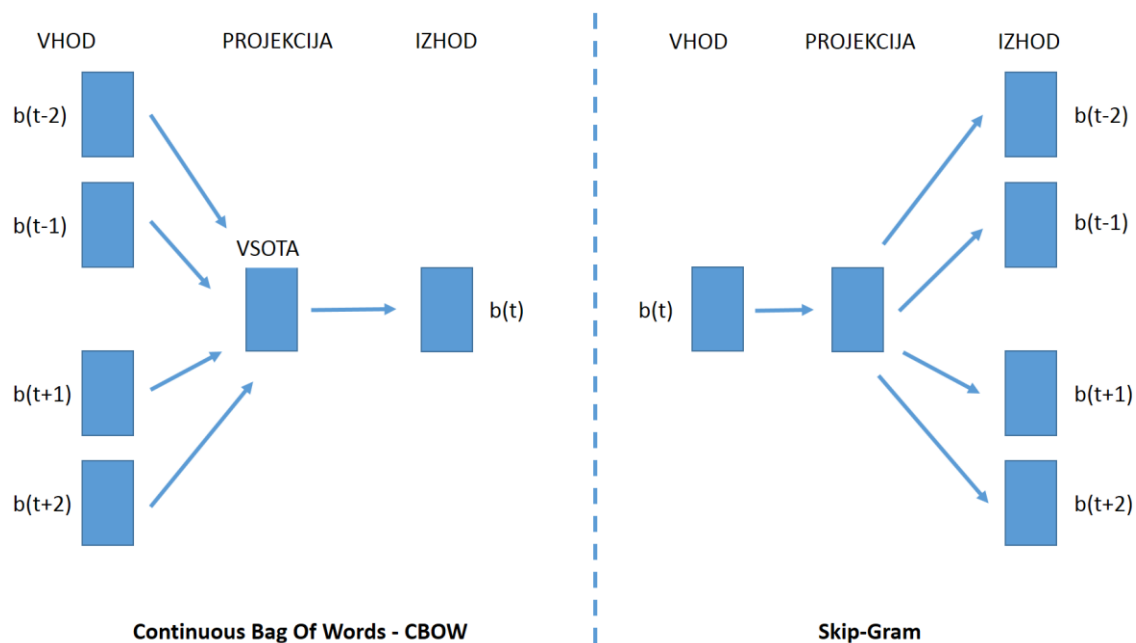
$$idf(b, B) = \log \frac{N}{|\{p \in B : b \in p\}|} \quad (2)$$

Kjer je  $N$  število povedi v besedilu in  $|\{p \in B : b \in p\}|$  število povedi v katerih se pojavi dana beseda. Vrednost IDF nam pove koliko informacij nosi beseda. Če bo torej beseda pogosta v dani povedi in se ne pojavlja pogosto v našem besedilu, bo imela visoko TF-IDF utež.

### 2.3.2 Word Embedding

Word embedding je tehnika, ki besede predstavi z vektorskim zapisom na tak način, da so si vektorji semantično podobnih besed bližje v vektorskem prostoru. Obstajajo različni algoritmi, ki to dosežejo. Dobra primera sta Word2Vec (Mikolov, Chen, Corrado & Dean, 2013) in Facebookov fastText. Ti algoritmi običajno delujejo na podlagi nevronske mreže naučene na določenem besedilu ali večji količini teksta. Dva modela, ki sta pogosto uporabljena za učenje nevronske mreže namenjene vektorizaciji sta Continuous bag of words in Skip-Gram. Slika 1 prikazuje delovanje omenjenih modelov.

Slika 1: Ilustracija delovanja modelov CBOW in Skip-Gram



Prerejeno po Mikolov & Sutskever (2013).

Model CBOW poskuša na podlagi konteksta, ki je podan s predhodnimi in sledečimi besedami, napovedati manjkajočo besedo. Obratno poskuša model Skip-Gram na podlagi dane besede sestaviti kontekst oziroma predhodne in sledeče besede. Oba modela je možno naučiti na danem vhodnem besedilu, s pomočjo katerega se model nauči združevati besede v kontekst na podlagi pravih in napačnih napovedi. Izhod takšnih modelov so besede (lahko tudi celotne povedi oziroma zaporedja besed) predstavljene z vektorji. Vektorje lahko uporabimo za različne matematične operacije. Pogosto uporabljen primer delovanja takšne vektorizacije so besede kralj, moški, ženska in kraljica. Enačbe (3), (4), (5) in (6) prikazujejo kako bi te besede izgledale, če bi jih predstavili z vektorji dolžine 3:

$$kralj = [5, 9, 5] \quad (3)$$

$$moški = [5, 2, 2] \quad (4)$$

$$\textit{zenska} = [3, 4, 3] \quad (5)$$

$$\textit{kraljica} = [3, 11, 6] \quad (6)$$

Zgornji vektorji omogočajo sledečo operacijo (7):

$$\textit{kralj} - \textit{moški} + \textit{zenska} = \textit{kraljica} \quad (7)$$

Vidimo tudi, da sta si vektorja besed *moški* in *zenska* v vektorskem prostoru bližje kot recimo *moški* in *kralj* ter *moški* in *kraljica*. Prav tako sta si bližja vektorja *kralj* in *kraljica* v primerjavi z razdaljo med vektorjema *kralj* in *zenska*. Besede so v vektorskem prostoru teh modelov seveda predstavljene z daljšimi vektorji. To dolžino lahko predhodno določimo in je običajno med 50 in 300 (Patel & Bhattacharyya, 2017).

### 2.3.3 ELMo in BERT

Embedding from Language Model je tehnika vektorizacije, ki upošteva ne samo semantiko besede ampak tudi sintakso, hkrati pa ima sposobnost razdvoumljanja večpomenskih besed (Singh & Shashi, 2019). Na isti ideji je zasnovan najnaprednejši model: Googlov BERT - Bidirectional Encoder Representations from Transformers (Devlin, Chang, Lee & Toutanova, 2018). BERT s pridom izkorišča dva načina, ki omogočata nadzorovano učenje brez človeško anotiranih podatkov. Prvi način je zakrivanje (angl. masking) določenih besed v povedi, ki od modela zahteva, da napove pravilno manjkajočo besedo v povedi. Ker je manjkajoča beseda znana, lahko modelu sporočimo, ali se je odločil prav ali ne. Tako se model izboljša pri povezovanju besed v isti kontekst. Med učenjem modela BERT so ustvarjalci skrili 15% besed (Devlin, Chang, Lee & Toutanova, 2018). Drugi način učenja brez anotiranih podatkov je napovedovanje naslednje povedi. Pri tem načinu so modelu predstavili dve povedi, od njega pa zahtevali, da se odloči, ali ena poved logično sledi drugi ali ne. S tem se model uči povezovanja povedi v kontekst. Vhodni podatki na katerih je naučen model BERT so besedila iz knjig (800 milijonov besed) in angleške Wikipedie (2500 milijonov besed). Velika uporabnost modela BERT pa je v dejstvu, da ga lahko prilagodimo glede na našo specifično nalogo ali domeno besedil. To storimo s postopkom imenovanim »fine-tuning«. V praksi to pomeni, da znanje modela nadgradimo z našimi ali sorodnimi podatki, kar bo zvišalo napovedno zmogljivost oziroma natančnost.

## 2.4 Analiza sentimenta ali analiza razpoloženja

Zelo razširjeno področje rudarjenja besedil je analiza sentimenta oziroma analiza razpoloženja (tudi rudarjenje mnenj). Ta nam omogoča, da ugotovimo, ali naše besedilo oziroma izbrane povedi izražajo določeno čustvo in razpoloženje. Običajno nas zanima, ali je v besedilu moč zaznati pozitivno ali negativno naravnost pisca oziroma sporočevalca. Nekatere analize se še bolj poglobijo in poskušajo zaznati določeno čustvo, kot so recimo

jeza, presenečenje, sreča itn. Uporabnost analize razpoloženja je zelo velika. Podjetja lahko tako poskušajo ugotoviti, kakšno je javno mnenje o novem produktu, investitorji lahko ugotavljajo, kakšno je splošno razpoloženje povedi v letnem poročilu podjetja, politike zanima, kakšen je odziv javnosti na njihov govor in še bi lahko našteval. Analiza sentimenta je ena od analiz, ki jih opravi za ugotavljanje povezave med dejavniki ESG in finančno uspešnostjo podjetij. Tehnik za analizo sentimenta je več. Najbolj osnovni pristop uporablja pristop vreče besed (angl. bag of words). Pri tej tehniki začnemo s seznamoma pozitivnih in negativnih besed in enostavno preštevamo, katere se v povedi pojavijo večkrat. Bolj napredne tehnike se osredotočajo na ustrezno vektorizacijo povedi in običajno vključujejo seznam označenih oziroma anotiranih povedi. Gre torej za nadzorovano učenje, kjer se model nauči katere povedi (predstavljene z vektorji) naj poveže z določenim razpoloženjem.

### **3 RAZISKAVA POVEZAVE MED DEJAVNIKI ESG IN FINANČNO USPEŠNOSTJO PODJETIJ**

#### **3.1 Opredelitev raziskovalnih vprašanj in metodologija**

Prvo in glavno raziskovalno vprašanje temelji na pregledu ugotovitev sorodnih študij in je:

#### **V1: Ali obstaja pozitivna povezava med ESG oceno in finančno uspešnostjo podjetja?**

Prvi del raziskovanja postavljenega vprašanja bo temeljil na metodi primerjave. 100 izbranih podjetij sem s tem namenom razdelil v skupino 50 največjih in 50 najmanjših po tržni kapitalizaciji. Obe skupini bom naprej primerjal po povprečnih letnih vrednostih kazalnikov izračunanih na podlagi finančnih izkazov in tržnih podatkov in nato še po merah ESG ocene. Predvidevam, da bo skupina največjih 50 dosegala boljše rezultate pri kazalnikih finančne uspešnosti kot tudi pri merah ESG ocene. Tudi drugi del raziskovanja prvega vprašanja temelji na metodi primerjave. Podjetja razdelim v novi dve skupini glede na povprečno ESG oceno. Skušal bom ugotoviti, ali se novi dve skupini pomembno razlikujeta po finančni uspešnosti izraženi z definiranimi kazalniki. Tretji del raziskovanja prvega vprašanja temelji na izračunu Pearsonovega korelacijskega koeficienta med ESG oceno in posameznimi merami ESG ocene ter finančnimi in tržnimi kazalniki.

Postavljam tudi drugo raziskovalno vprašanje, ki temelji na ugotovitvah študij, da obstajajo pomembne razlike pri dojemanju dejavnikov ESG glede na gospodarsko panogo ter da so določene dejavnosti izpostavljene večjim ESG tveganjem.

#### **V2: Ali obstajajo pomembne razlike v ESG ocenah podjetij glede na panogo, ki ji pripadajo?**

Pri drugem raziskovalnem vprašanju postavljam še dodatno podvprašanje:

### **V3: Ali podjetja, ki delujejo v panogah z večjim ESG tveganjem, dosegajo povprečno nižje ESG ocene?**

Pri raziskovanju drugega in tretjega vprašanja podjetja razdelim po panogah in se ponovno opiram na metodo primerjave. Iskal bom razlike med panogami glede na ESG oceno in posamezno mero ESG ocene. Predvidevam, da bom lahko opazil jasne razlike v povprečnih vrednostih posameznih mer glede na panogo. Dodatno pričakujem, da bo pri podjetjih, ki pripadajo panogi z višjim ESG tveganjem, moč opaziti višje število in odstotek ESG povedi, a hkrati nižje ESG razpoloženje v primerjavi s panogami z nižjim ESG tveganjem.

Podpoglavja v poglavju 3 si sledijo v istem zaporedju, kakor je izgledal pristop k raziskovanju postavljenih domnev. Najprej sem izbral seznam podjetij, katerih letna poročila bodo vključena v analizo. Nato sem določil kazalnike, s katerimi izrazim finančno uspešnost podjetij. Sledila je definicija ESG mer, ki bodo uporabljene za izračun ESG ocene podjetja. Pred empiričnim preverjanjem povezave med ESG in finančno uspešnostjo sem vhodne podatke pretvoril v računalniku razumljivo obliko s pomočjo različnih računalniških metod predprocesiranja. Temu je sledil izračun ESG mer in ESG ocene za vsa podjetja. Zadnji del analize je bila uporaba primerjalnih in statističnih metod za preverjanje povezave med ESG oceno in finančno uspešnostjo ter za ugotavljanje razlik v ESG ocenah med gospodarskimi panogami.

#### **3.2 Izbira podjetij**

Osrednji del raziskave je s pomočjo metod za rudarjenje besedil analizirati letna poročila podjetij iz skupine S&P 500. Za to skupino sem se odločil, ker so ta podjetja vsakoletno dolžna predložiti 10K poročila. Oblika 10K poročil omogoča, da z uporabo računalniških metod izvlečemo tiste dele besedila, ki nas zanimajo za analizo. Za to skupino sem se odločil tudi zaradi dobrega poznavanja oblike 10K poročil iz projektov, ki jih omenjam v uvodu magistrskega dela. 10K poročila podjetij je možno prenesti v html obliki preko sistema EDGAR. Pri letnih poročilih sem se omejil na zadnjih 5 finančnih let, ki so na voljo v sistemu (2017, 2018, 2019, 2020, 2021). Letna poročila so predložena ob koncu finančnega leta, ki pa se ne nujno sklada s koncem koledarskega leta. Vsi na leta vezani podatki podjetij in primerjave v magistrskem delu se nanašajo na finančna leta podjetij in ne na koledarska leta. Seznam sem omejil glede na povprečno tržno kapitalizacijo največjih 50 in najmanjših 50 podjetij iz skupine. To omogoča dodatno primerjavo med večjimi in manjšimi podjetji in iskanje razlik med ESG ocenami ter finančno uspešnostjo. Prav tako sem izbor omejil samo na podjetja, ki so vseh 5 izbranih let v indeksu S&P 500. Velikost podjetja sem kot že omenjeno določil na podlagi povprečne tržne kapitalizacije v analiziranih petih letih. Podatki o tržnih kapitalizacijah so bili pridobljeni na spletni strani: <https://companiesmarketcap.com/>. Končni seznam letnih poročil tako vsebuje 5 poročil za vsako podjetje, kar je skupno 500 poročil. Spodnja tabela 2 prikazuje vsa analizirana podjetja razvrščena po povprečni tržni kapitalizaciji.

*Tabela 2: Skupini največjih 50 in najmanjših 50 podjetij glede na povprečno tržno kapitalizacijo od leta 2017-2021*

Največjih 50		Najmanjših 50	
1. Apple	26. Comcast	1. Quanta Services	26. PerkinElmer
2. Microsoft	27. Costco Wholesale	2. Assurant	27. Invesco
3. Amazon	28. Oracle	3. Robert Half	28. Jacobs
4. Alphabet	29. Pepsi	4. Sealed Air	29. Nielsen
5. Meta	30. Citigroup	5. PVH	30. Advance Auto Parts
6. Berkshire Hathaway	31. Boeing	6. Kimco Realty	31. AES
7. Johnson & Johnson	32. Nvidia	7. Ralph Lauren	32. Henry Schein
8. JPMorgan	33. AbbVie	8. News Corporation	33. Qorvo
9. Visa	34. Nike	9. Under Armour	34. Howmet Aerospace
10. Walmart	35. McDonald's	10. Zions Bancorporation	35. Comerica Incorporated
11. Exxon Mobil	36. Phillip Morris	11. Interpublic Group	36. Marathon Oil
12. Bank of America	37. Adobe	12. NRG Energy	37. Universal Health Services
13. Procter & Gamble	38. Medtronic	13. Allegion	38. Tractor Supply Company
14. UnitedHealth Group	39. Paypal	14. BorgWarner	39. DaVita
15. AT&T	40. Amgen	15. PulteGroup	40. CH Robinson Worldwide
16. Verizon Communications	41. Netflix	16. CF Industries Holding	41. FMC
17. Wells Fargo	42. IBM	17. Pentair	42. Bath & Body Works
18. Mastercard Incorporated	43. Abbott Laboratories	18. Iron Mountains	43. Hasbro
19. Home Depot	44. Eli Lilly	19. Juniper Networks	44. Eastman Chemical
20. Pfizer	45. Broadcom	20. Snap-on Incorporated	45. Textron
21. Walt Disney	46. Honeywell	21. F5	46. Masco
22. Intel	47. Salesforce	22. Coterra Energy	47. J B Hunt Transport Services
23. Chevron	48. Accenture	23. Mosaic	48. Dentsply Sirona
24. Coca-Cola	49. Bristol-Myers	24. Avery Denison	49. United Rentals
25. Merck & Co	50. 3M	25. Globe Life	50. International Flavors & Fragrances

*Vir: lastno delo, izračunano na podlagi podatkov pridobljenih iz <https://companiesmarketcap.com>.*

### 3.3 Kazalniki finančne uspešnosti

V analizo sem vključil več kazalnikov finančne uspešnosti, ki omogočajo primerjavo ESG ocene z določenim vidikom uspešnosti. Kazalniki so razdeljeni v dve skupini glede na podatke s katerimi so bili izračunani (finančni izkazi in tržne vrednosti). Z izbranimi kazalniki želim karseda dobro izraziti finančno uspešnost podjetij in tveganje. Izbor kazalnikov temelji tudi na pregledu sorodne literature. Konkretno o najbolj pogosto uporabljenih in poročanih kazalnikih finančne uspešnosti pišem v poglavju 1.2. Po pregledu sorodnih študij zaključim, da finančno uspešnost in tveganje dobro povzamejo sledeči kazalniki: prihodki, bruto marža, neto marža, prosti denarni tok, dobiček pred obrestmi in davki (angl. earnings before interest and taxes, v nadaljevanju EBIT), donosnost sredstev, donosnost kapitala in donosnost delnic ter standardni odklon in varianca donosnosti delnic. V magistrskem delu med seboj primerjam podjetja različnih velikosti skozi 5 analiziranih let, zato vsi kazalniki finančne uspešnosti predstavljajo relativne vrednosti. Končna seznama izbranih kazalnikov sta sledeča:

#### Kazalniki na podlagi finančnih izkazov:

- **rast prihodkov** (angl. revenue growth) izračunana po enačbi (8):

$$\text{rast prihodkov} = \left( \left( \frac{\text{prihodki v letu } n}{\text{prihodki v letu } n-1} \right) - 1 \right) * 100\% \quad (8)$$

- **bruto marža** (angl. gross margin) izračunana po enačbi (9):

$$\text{bruto marža} = \left( \frac{\text{bruto dobiček}}{\text{prihodki}} \right) * 100\% \quad (9)$$

- **neto marža** (angl. profit margin) izračunana po enačbi (10):

$$\text{neto marža} = \left( \frac{\text{čisti dobiček}}{\text{prihodki}} \right) * 100\% \quad (10)$$

- **marža prostega denarnega toka** (angl. free cash flow margin) izračunana po enačbi (11):

$$\text{marža prostega denarnega toka} = \left( \frac{\text{prosti denarni tok}}{\text{prihodki}} \right) * 100\% \quad (11)$$

kjer je prosti denarni tok izračunan po enačbi (12):

$$\text{prosti denarni tok} = \text{prihodki od prodaje} - (\text{stroški poslovanja} + \text{davki}) - \text{potrebne investicije v dolgoročna sredstva in v neto obratni kapital} \quad (12)$$

- **marža iz poslovanja** izračunana po enačbi (13):

$$\text{marža iz poslovanja} = \left( \frac{EBIT}{\text{prihodki}} \right) * 100\% \quad (13)$$

- **ROA** izračunan po enačbi (14):

$$ROA = \frac{\text{čisti dobiček}}{\text{povprečna skupna sredstva}} \quad (14)$$

- **ROE** izračunan po enačbi (15):

$$ROE = \frac{\text{čisti dobiček}}{\text{povprečno stanje lastniškega kapitala}} \quad (15)$$

### Kazalniki na podlagi tržnih podatkov:

- **povprečna letna donosnost delnic** glede na dnevne vrednosti donosnosti izračunane po enačbi (16):

$$\text{donosnost delnice} = \left( \left( \frac{\text{vrednost delnice na dan } n}{\text{vrednost delnice na dan } n-1} \right) - 1 \right) * 100\% \quad (16)$$

- **standardni odklon donosnosti delnic** izračunan po enačbi (17):

$$\text{standardni odklon } \sigma = \sqrt{\frac{(x_1 - \bar{x}) + (x_2 - \bar{x}) + \dots + (x_n - \bar{x})}{N}} \quad (17)$$

kjer je  $x_n$  donosnost na dan  $n$ ,  $\bar{x}$  povprečna letna donosnost in  $N$  število vrednosti,

- **varianca donosnosti delnic**, ki je enaka  $\sigma^2$ .

Za vsakega od naštetih kazalnikov sem potreboval zanesljiv vir, ki je omogočal zbiranje natančnih podatkov za vsa analizirana podjetja in vseh 5 vključenih let. Finančni izkazi z vsemi potrebnimi podatki na katerih temelji prva skupina kazalnikov so bili pridobljeni na spletni strani: <https://stockanalysis.com/stocks/>. Tržni podatki so bili pridobljeni s pomočjo Yahoo! Finance aplikacijskega programskega vmesnika (angl. application programming interface, v nadaljevanju API) s katerim lahko v Pythonu lažje komuniciramo s pomočjo knjižnice yfinance.



### 3.4 ESG ocena podjetij

ESG ocena temelji na analizi besedil letnih poročil podjetij. Metode analize teksta omogočajo izračun različnih mer in kazalnikov. V ESG oceno vključim tiste mere, ki celovito zajamejo uspešnost podjetja, ki jo lahko zaznamo in merimo s pomočjo analize teksta.

- **Število ESG povedi v poročilu.** Pri tej meri predpostavljam, da je višje število z vidika ESG uspešnosti boljše, saj lahko več povedi nosi večjo količino informacij, kar je za bralca, ki ga zanimajo ESG aktivnosti podjetja, pozitivno.
- **Odstotek ESG povedi v poročilu.** Dolžine letnih poročil merjene v številu povedi podjetij se razlikujejo. Z vključitvijo mere »odstotek ESG povedi« lahko merim, kolikšen del poročila podjetje posveti ESG temam. Predpostavim, da je višji odstotek z vidika ESG uspešnosti boljši.
- **Povprečni sentiment vseh povedi poročila.** Povprečni sentiment celotnega poročila predstavlja meritev splošnega razpoloženja v poročilu. Ta mera ni neposredno povezana z ESG uspešnostjo, vseeno pa jo uporabim za boljše razumevanja ESG razpoloženja, ki je merjena z zadnjima dvema merama na tem seznamu. Višje razpoloženje je z vidika ESG uspešnosti boljše.
- **Povprečni sentiment celotnega poročila brez nevtralnih povedi poročila.** V večini povedi letnih poročil ni moč zaznati prevladujočega pozitivnega ali negativnega razpoloženja. Da lahko bolje zaznam kontrast med negativnimi in pozitivnimi povedmi v poročilu, nevtralne povedi pri tej meri izključim pri računanju povprečnega sentimenta poročila. Enako kot pri prejšnji meri je višje razpoloženje boljše.
- **Sentiment ESG povedi v poročilu.** Ta mera zajema razpoloženje samo v tistih povedih, ki se nanašajo na ESG teme. Višje razpoloženje je boljše z vidika ESG uspešnosti.
- **Sentiment ESG povedi brez nevtralnih povedi poročila.** Enako kot pri razpoloženju v vseh povedih poročila tudi pri ESG sentimentu izračunam razpoloženje samo na podlagi povedi, ki so pozitivne ali negativne. Višje razpoloženje doprinese k višji skupni ESG oceni.

ESG oceno za vsako podjetje pridobim s sledečim postopkom:

1. Vsa podjetja razvrstim od 1 do 100 glede na izračunano povprečno vrednost posamezne mere v določenem finančnem letu.
2. Glede na uvrstitev na lestvici dodelim podjetju ustrezno število točk. Prvo ali najboljšo podjetje dobi 100 točk, drugo 99 in tako naprej do 1.
3. Postopek ponovim za vseh 6 zgoraj naštetih mer.
4. Vseh 6 uvrstitev seštejem in vsoto uporabim kot ESG oceno podjetja.

## 3.5 Priprava podatkov

Ključni del analize je ustrezna priprava podatkov, ki bo omogočala uporabo računalniških metod za rudarjenje besedil. Spodnji koraki prikazujejo pridobivanje in obdelavo podatkov, kot so si dejansko sledili pri pripravi za analizo.

### 3.5.1 Prenos letnih poročil izbranih 100 podjetij

Prvi korak je prenos vseh 500 letnih poročil. Večina poročil je bila na voljo preko sistema EDGAR dostopnega na <https://www.sec.gov/edgar/searchedgar/companysearch.html>. Nekatera poročila, ki niso bila najdena z iskanjem na zgornji povezavi, so bila na voljo na povezavi <https://sec.report/>. Poročila so bila shranjena v obliki html, ki jo je možno natančno razčleniti s pomočjo knjižnic programskega jezika Python (sam uporabim knjižnico BeautifulSoup).

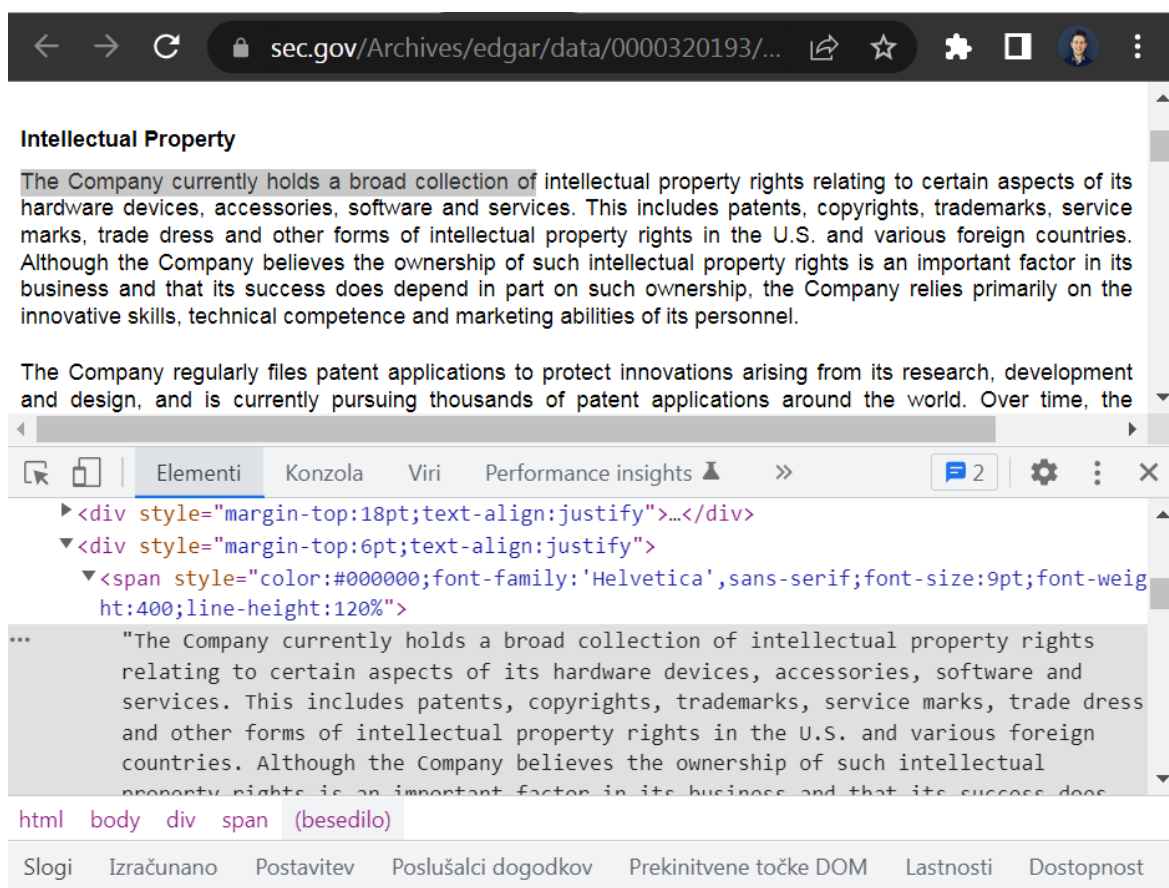
### 3.5.2 Ekstrakcija povedi iz poročil

Postopek in izzive ekstrakcije povedi iz danega besedila opišem v poglavju 2.1.1. Letno poročilo v html obliki ne omogoča enostavne razčlenbe v povedi. Sestavljeno je iz različnih grafičnih elementov, tabel in razpredelnic, naštevanj in seznamov. To vse oteži detekcijo povedi, ki se praviloma začnejo z veliko začetnico in končajo s piko. Lepše kot je strukturirana vhodna oblika podatkov, bolj natančno lahko NLTK-jev algoritem za ekstrakcijo povedi opravi svojo nalogo.

Celotna ESG analiza magistrskega dela temelji na povedih, zato je karseda natančna detekcija povedi ključna za dobre rezultate. Pred uporabo algoritma za ekstrakcijo povedi, sem tako letno poročil razčlenil s pomočjo tehnik html razčlenjevanja (angl. parsing). Spletna stran napisana v jeziku HTML je razdeljena na elemente. Elementi imajo svoje identifikatorje in določene stilistične attribute. S pomočjo teh lahko iz spletne strani pridobimo samo vsebino tistih elementov, ki jih smatramo kot relevantne za analizo. Za namen analize v magistrskem delu so me zanimali samo tisti elementi, ki vsebujejo povedi. To izključi tabele, ki jih je praktično nemogoče razčleniti v povedi, naslove in podnaslove, kazala, sezname, naštevanja ipd. To dosežem z definiranjem določenih stilističnih pravil, ki jih mora html element izpolnjevati, da je vključen v analizo. Slika 2 prikazuje strukturo html strani in attribute elementov.

Ko je ta korak zaključen, je besedilo uporabljeno kot vhodni podatek v NLTK funkcijo sentence tokenizer, ki vrne seznam vseh povedi v besedilu.

Slika 2: HTML elementi in pripadajoči slogovni atributi



Vir: lastno delo, posnetek delovnega okna programa Google Chrome.

### 3.5.3 Ekstrakcija ESG povedi

Drugi ključni korak priprave podatkov je ekstrakcija ESG povedi. V tem koraku iz seznama vseh povedi pridobim povedi, ki se navezujejo na vsaj enega od vidikov ESG. Za reševanje tega problema uporabim strojno učenje in seznam s strani študentov anotiranih povedi (študenti Ekonomske fakultete, 2021). Seznam vsebuje 2629 povedi, ki so označene kot »ESG« ali »NI ESG«. Z označenimi podatki lahko uporabimo enega od algoritmov nadzorovanega učenja, da naučimo napovedni model, ki se bo znal odločiti, ali se poved navezuje na ESG teme ali ne. Pred tem pa je najprej potrebno podatke pretvoriti v računalniku razumljivo obliko. Pri tem sledim korakom predprocesiranja opisanim v poglavju 2.1.2. Po obdelavi povedi uporabim algoritem TFIDF, ki je opisan v poglavju 2.2.1 in pretvori povedi v vektorski zapis. Po pretvorbi uporabim povedi v vektorskem zapisu in pripadajoče »ESG« in »NI ESG« oznake kot vhodna podatka v algoritmu za učenje napovednega modela. Za to nalogo uporabim metodo podpornih vektorjev (angl. support-vector machine, v nadaljevanju SVM). Naučeni model sem testiral tako, da sem anotirane povedi razdelil v dve skupini. Prva skupina je namenjena učenju (80% povedi), druga skupina pa evalvaciji natančnosti (merjena z odstotkom pravih napovedi) algoritma (20%

povedi). Rezultat evalvacije je več kot 70% pravih napovedi. Preizkusil sem tudi nekatere druge algoritme nadzorovanega učenja, a ti niso dosegli višjega odstotka pravih napovedi (povprečno okoli 65%). Iz tega razloga izberem SVM tudi za končni model, ki ga uporabim za klasifikacijo povedi iz letnih poročil. Tabela 3 prikazuje napovedi modela za 20 povedi, ki sem jih sestavil sam. Prvih 10 povedi ni povezanih z ESG temami, medtem ko drugih 10 je.

*Tabela 3: Prikaz napovedi ESG napovednega modela*

Poved	Dejanska vrednost	Napovedana vrednost
Spring is finally here.	NI ESG	NI ESG
Winter is going to be very cold.	NI ESG	ESG
Game of thrones was a great series.	NI ESG	NI ESG
I can't wait for the Lord of the rings series to come out.	NI ESG	NI ESG
Easter is just around the corner, here goes my diet.	NI ESG	ESG
The sun is shining.	NI ESG	NI ESG
Grass is green and sky is blue.	NI ESG	NI ESG
Revenue will go up next year.	NI ESG	NI ESG
My car broke down and I had to go to the mechanic.	NI ESG	ESG
We're going to the movies tomorrow evening.	NI ESG	ESG
We ensure our employees have a friendly working environment.	ESG	ESG
Sustainability is of the top importance for our company.	ESG	ESG
We upgraded our air pollution control devices last month.	ESG	NI ESG
We strictly prohibit any kind of work place environment.	ESG	ESG
We provide monthly courses for our employees.	ESG	ESG
Leadership and independence are highly encouraged.	ESG	NI ESG
Our business model is very environmentally friendly	ESG	ESG
Ethnicity plays no role in our hiring process.	ESG	ESG
We employ the same number of men and women.	ESG	ESG
We reduced our emissions by 20% last year and will continue to reduce them further.	ESG	ESG

*Vir: lastno delo.*

#### 3.5.4 Analiza sentimenta ali razpoloženja

Osrednji del analize povedi je analiza razpoloženja. To nam pove, ali je v povedi moč zaznati pozitivno ali negativno razpoloženje sporočevalca. Na podlagi te analize kasneje ugotavljam, ali obstaja povezava med prevladujočim razpoloženjem v poročilu in finančno uspešnostjo podjetja. Prav tako me je zanimalo, kakšne so razlike v razpoloženju med največjimi in najmanjšimi podjetji v skupini S&P 500. Nekaj tehnik za analizo razpoloženja predstavim že v poglavju 2.4. V tem delu za analizo razpoloženja uporabim Googlov model BERT. Izdaja tega predstavlja velik korak naprej na področju NLP in poveča zmožnosti in učinkovitost reševanja NLP nalog (Alberti, Lee & Collins, 2019). Največja ovira pri izdelavi

modelov za reševanje NLP nalog je pomanjkanje anotiranih oziroma označenih podatkov. V prejšnjem poglavju sem imel na voljo več kot 2600 anotiranih povedi, ki jih je napovedni model uporabil za učenje. Če bi bilo anotiranih povedi še več, bi lahko bil model še boljši. Najnaprednejše metode za učenje napovednih modelov uporabljajo v naprej naučene modele (angl. pretrained), ki so naučeni na veliki količini podatkov. Tak model lahko kasneje dodatno prilagodimo za specifične zahteve našega problema (z že omenjenim postopkom fine-tuning). V tem magistrskem delu so me bolj kot splošna besedila zanimala finančna poročila. Za analizo sentimenta sem tako uporabil model BERT, ki je bil dodatno izboljššan na besedilih iz finančnih novic. Uporabljen model je na voljo na: <https://huggingface.co/mrm8488/distilroberta-finetuned-financial-news-sentiment-analysis>. Za vsako poved model vrne eno izmed vrednosti -1, 0 in 1, kjer -1 predstavlja negativno razpoloženje, 0 nevtralnno in 1 pozitivno. Spodnja tabela 4 prikazuje delovanje modela na 15 povedih, ki sem jih sestavil sam. Prvih 5 je negativnih, drugih 5 nevtralnih in zadnjih 5 pozitivnih. Ta model uporabim za napovedovanje razpoloženja vseh povedi v poročilu in ESG povedi v poročilu. Natančnost (odstotek pravilnih napovedi) modela BERT za analizo sentimenta je višja od 75% (Karimi, Rossi & Prati, 2020).

*Tabela 4: Prikaz napovedi modela BERT za klasifikacijo razpoloženja*

Poved	Dejanska vrednost	Napovedana vrednost
Our debt has increased significantly this year.	Negativno	Pozitivno
Revenue will drop as a result of the Covid crisis.	Negativno	Negativno
Salaries of our employees are lower compared to our competitors.	Negativno	Negativno
We lost two production buildings in a fire in July.	Negativno	Negativno
We were unsuccessful in reducing our carbon footprint.	Negativno	Pozitivno
Business operations are proceeding normally.	Nevtralnno	Nevtralnno
Our production buildings are painted green.	Nevtralnno	Nevtralnno
We require our employees to wear professional clothing at our work place.	Nevtralnno	Nevtralnno
Coffee breaks are at 11 am.	Nevtralnno	Nevtralnno
We also serve vegan meals.	Nevtralnno	Nevtralnno
Our debt has decreased significantly this year.	Pozitivno	Pozitivno
Revenue will go up as a result of successfully dealing with the Covid crisis.	Pozitivno	Pozitivno
Salaries of our employees are higher compared to our competitors.	Pozitivno	Pozitivno
We upgraded fire detection systems at our production buildings.	Pozitivno	Pozitivno
We were successful in reducing our carbon footprint.	Pozitivno	Pozitivno

*Vir: lastno delo.*

### 3.5.5 Izračun kazalnikov finančne uspešnosti

Definirane kazalnike izračunam na podlagi pridobljenih finančnih izkazov in podatkih o cenah delnic podjetij. Vse vrednosti so shranjene v excelovih datotekah, s katerimi lahko zelo učinkovito upravljamo s pomočjo Python knjižnice Pandas.

## 3.6 Metodologija izračuna povprečne korelacije

Med preverjanjem povezave med ESG ocenami, njenimi merami in finančno uspešnostjo se odločim tudi za izračun korelacijskih koeficientov. Ker želim izračunati korelacijo med dvema zveznima spremenljivkama, se odločim za uporabo Pearsonovega koeficienta korelacije. Korelacijo izračunam za vsako podjetje posebej na podlagi vrednosti spremenljivk skozi pet analiziranih finančnih letih. V naslednjem koraku želim to korelacijo oceniti za celoten seznam izbranih podjetij. Najbolj preprosta rešitev tega problema je uporaba aritmetične sredine, kjer vse korelacije seštejemo in nato vrednost delimo s številom prisotnih koeficientov. Težava, ki se pojavi pri uporabi te metode, je pristranskost (angl. bias), do katere pride zaradi višjega koeficienta asimetrije, ko je absolutna vrednost korelacije večja od 0 (Silver & Dunlap, 1987; Corey, Dunlap & Burke, 1998). Posledično izračunana povprečna vrednost korelacije pogosto prenizko oceni korelacijo v celotni populaciji. Silver in Dunlap (1997) ter Corey, Dunlap in Burke (1998) kot možno boljše rešitev za izračun povprečne korelacije predlagajo uporabo Fisherjeve z transformacije (Fisher, 1915). To priporočilo še posebej velja za manjše vzorce. Po tej metodi računanja povprečja korelacijskih koeficientov najprej pridobljene koeficiente  $r$  pretvorimo v Fisherjevo  $z'$  vrednost po enačbi (18).

$$z' = 0.5 \ln \frac{1+r}{1-r} \quad (18)$$

Nato izračunamo povprečno vrednost vseh pridobljenih vrednosti  $z'$ . Izračunano povprečno vrednost na koncu pretvorimo nazaj v vrednost  $r$  po enačbi (19).

$$r = \frac{e^{2z'} - 1}{e^{2z'} + 1} \quad (19)$$

Prednost te metode je manjši koeficient asimetrije porazdelitve  $r$  vrednosti in s tem višja možnost za bolj točno oceno korelacije populacije. Iz tega razloga tudi sam uporabim predstavljeno metodo za izračun povprečnih korelacij.

Ratner (2009) predstavi splošno sprejete smernice za interpretacijo moči korelacije, ki jim sledim tudi sam v tem magistrskem delu.

## 4 REZULTATI

### 4.1 Primerjava največjih 50 in najmanjših 50 podjetij iz skupine S&P 500

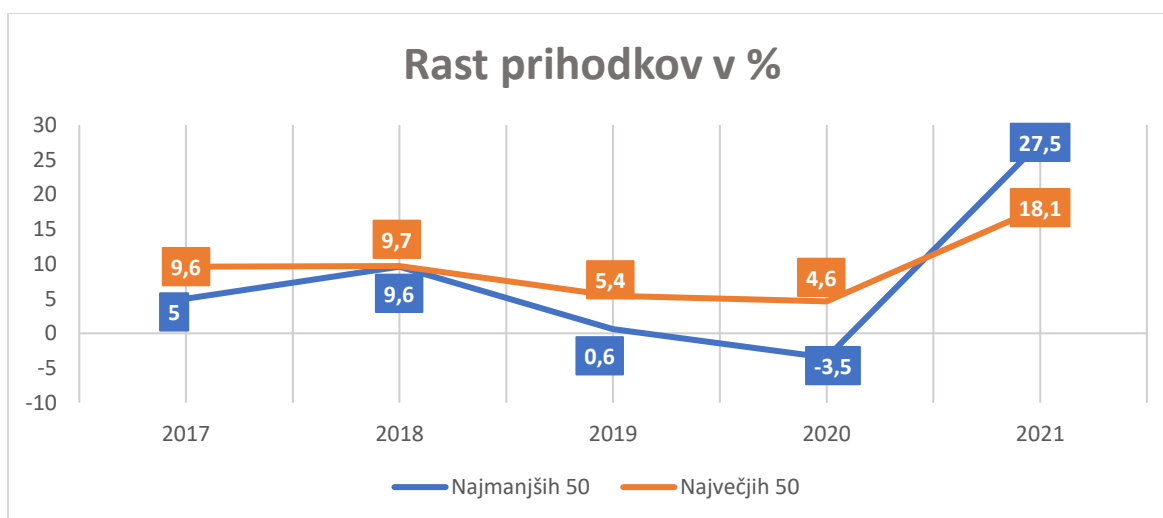
V prvem sklopu rezultatov primerjam največjih 50 podjetij z najmanjšimi 50 glede na povprečno tržno kapitalizacijo. Pri primerjavi finančnih in tržnih kazalnikov sem pričakoval jasno razliko v korist skupine največjih 50. Zanimalo me je predvsem, ali bo možno takšno razliko opaziti tudi pri rezultatih analize teksta.

#### 4.1.1 Primerjava finančnih kazalnikov

Spodnje slike 3-9 prikazujejo vse v analizo vključene finančne kazalnike, ki so predstavljeni v poglavju 3.3. Prikazane so povprečne vrednosti kazalnika združene glede na skupino podjetij ter razdeljene po finančnih letih.

Slika 3 prikazuje rast prihodkov obeh skupin. Opazimo lahko padec od leta 2018 do 2020 in nato rast do zadnjega analiziranega leta, v katerem skupina najmanjših 50 podjetij celo preseže skupino največjih 50. Vsa ostala leta dosegata skupina največjih boljše rezultate.

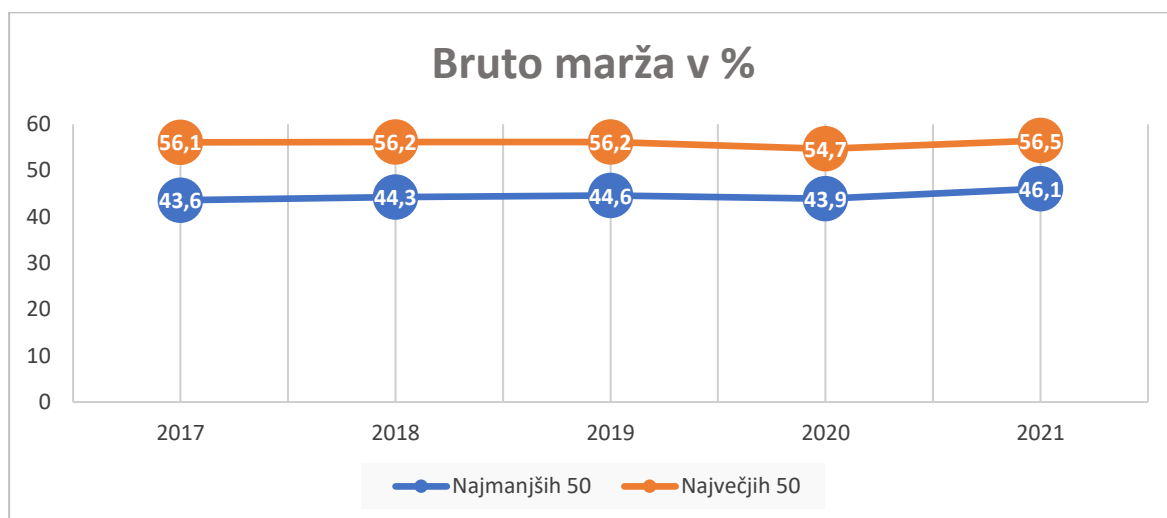
Slika 3: Primerjava rasti prihodkov največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Slika 4 omogoča primerjavo obeh skupin po kazalniku bruto marža. Največjih 50 z boljšimi rezultati skozi vsa leta.

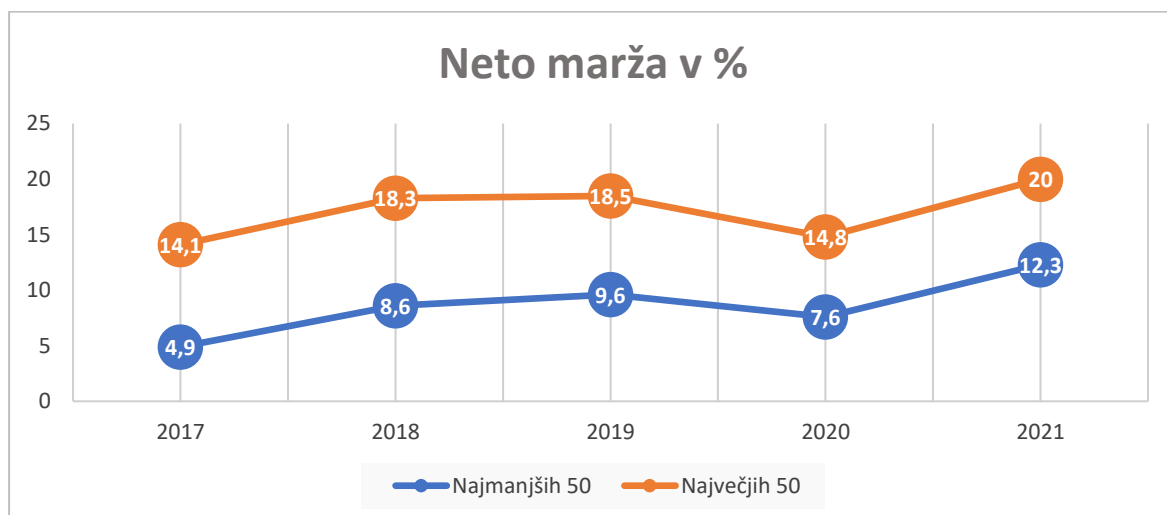
Slika 4: Primerjava bruto marže največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Kazalnik neto marža je prikazan na sliki 5. Podobno kot pri rasti prihodkov lahko opazimo padec od leta 2018 do 2020 in nato rast v letu 2021. Trend je zelo podoben za obe skupini. Tudi tukaj dosega skupina največjih 50 podjetij boljše rezultate.

Slika 5: Primerjava neto marže največjih 50 in najmanjših 50 podjetij

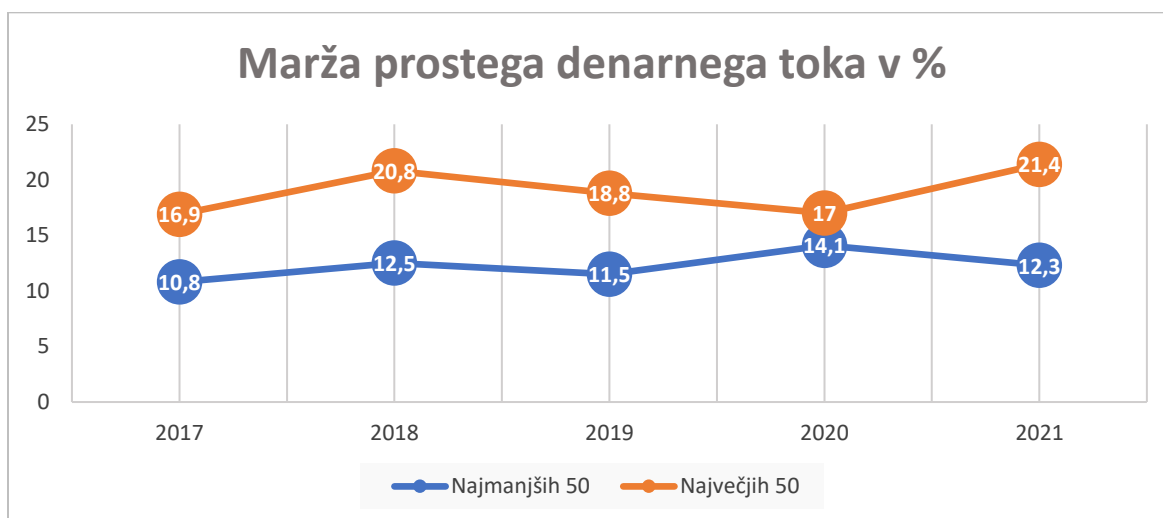


Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Slika 6 primerja marže prostega denarnega toka obeh skupin. Skupina največjih 50 z višjimi vrednostmi skozi vsa analizirana finančna leta.



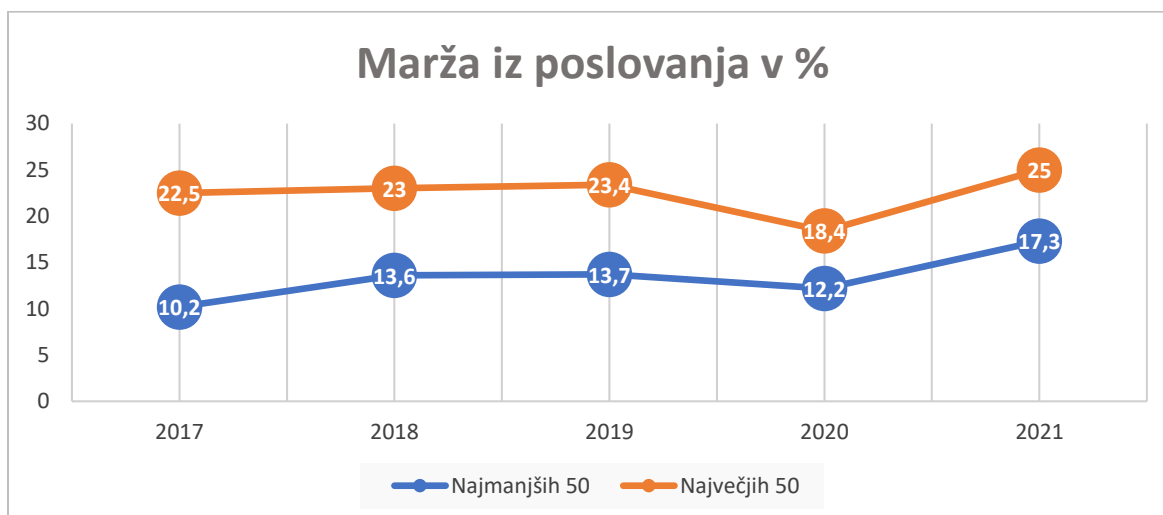
Slika 6: Primerjava marže prostega denarnega toka največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Kazalnik marža iz poslovanja je prikazan na sliki 7. Tudi po tej primerjavi dosega skupina največjih 50 podjetij boljše rezultate, prav tako pa je ponovno moč opaziti padec vrednosti v letu 2020.

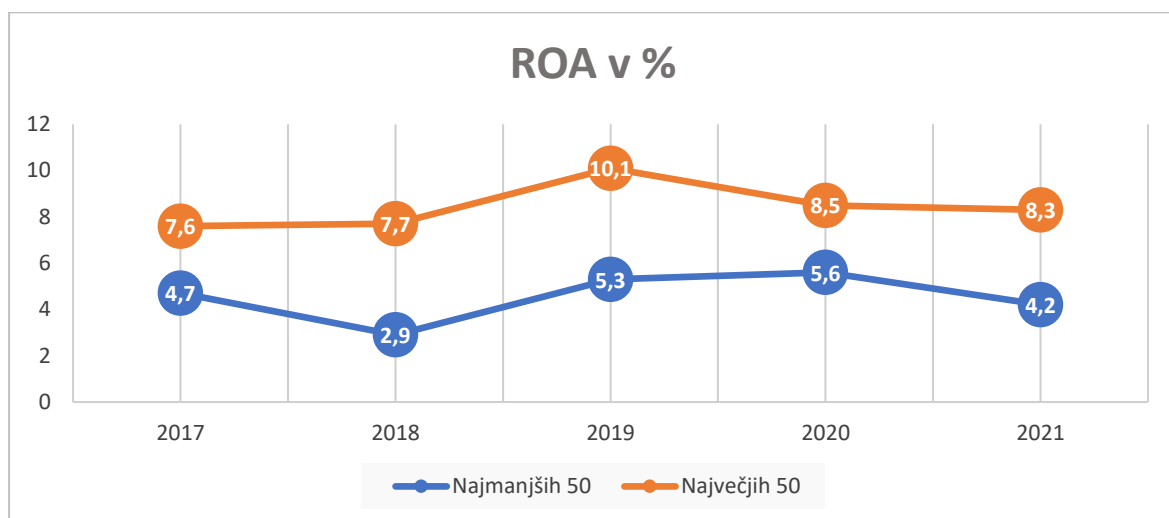
Slika 7: Primerjava marže iz poslovanja največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Slika 8 primerja skupini po kazalniku ROA. Skupina najmanjših 50 podjetij dosega slabše rezultate vseh 5 let.

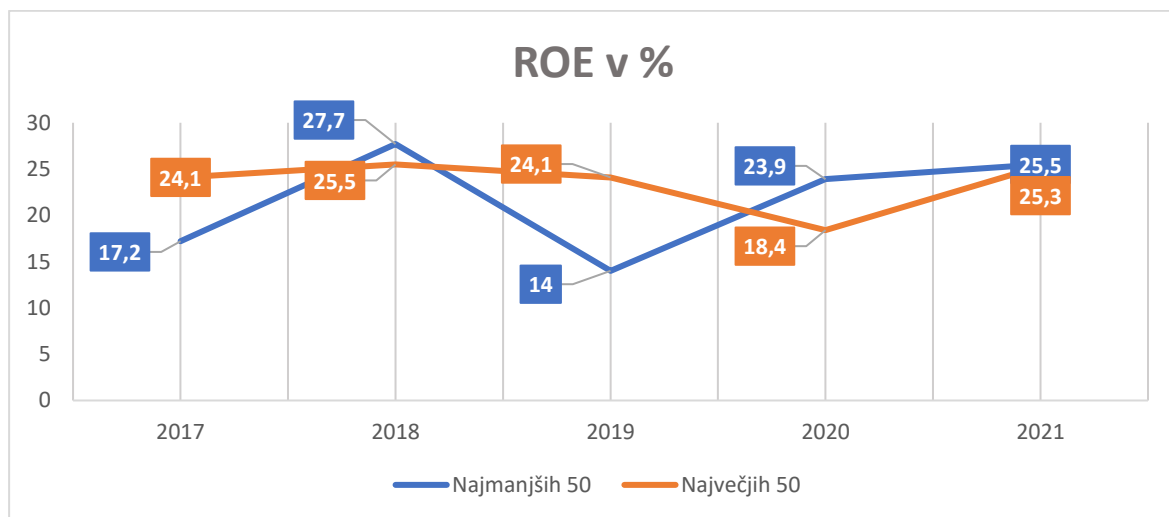
Slika 8: Primerjava ROA največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Zadnji kazalnik v tem sklopu je ROE, ki je prikazan na sliki 9. To je edini kazalnik, kjer skupina najmanjših 50 podjetij doseže boljši rezultat v več kot enem analiziranem letu. Kljub temu pa primerjava skupne povprečne vrednosti kazalnika za vseh 5 let, kar je razvidno na sliki 10, še vedno potrjuje prednost skupine največjih 50.

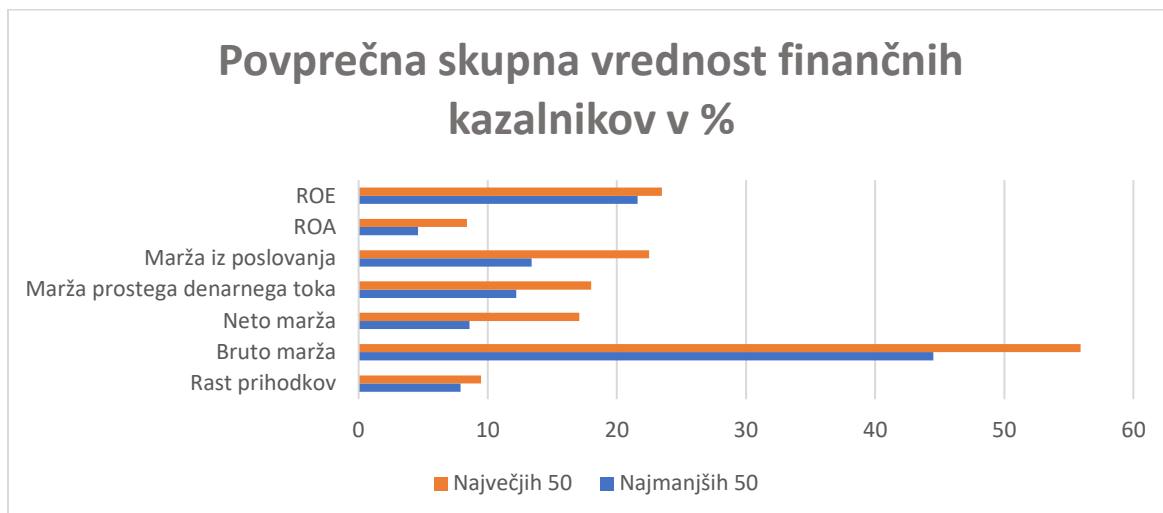
Slika 9: Primerjava ROE največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Primerjava finančnih kazalnikov da pričakovane rezultate. Skupina največjih 50 podjetij po vseh kazalnikih dosega boljše rezultate in to za praktično vsa merjena leta. Zanimivo je dejstvo, da je trend skozi 5 let zelo podoben za obe skupini pri več kazalnikih (rast prihodkov, bruto marža, neto marža in marža iz poslovanja).

Slika 10: Primerjava povprečne skupne vrednosti finančnih kazalnikov največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

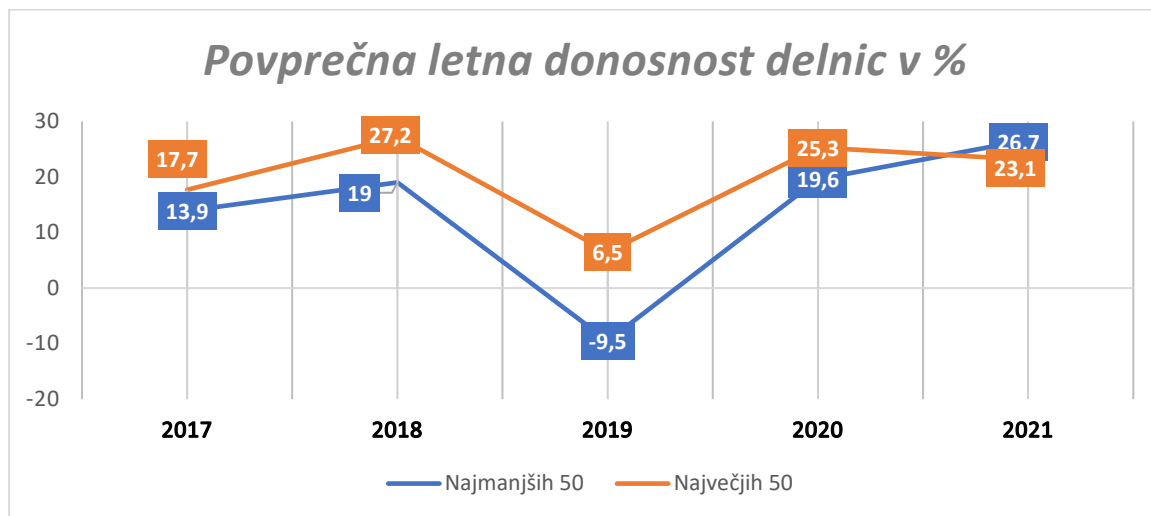
Z izjemo ROA lahko pri prav vseh kazalnikih opazimo padec obeh skupin od leta 2018 do 2020 in rast od leta 2020 do 2021 (izjema marža prostega denarnega toka pri skupini najmanjših 50). Kot možen razlog za ta pojav lahko navedem krizo Covid 19, ki je očitno podjetja obeh skupin prizadela v enaki ali podobni meri.

#### 4.1.2 Primerjava tržnih kazalnikov

Slike 11-13 omogočajo primerjavo kazalnikov pridobljenih na podlagi tržnih podatkov. Slika 11 prikazuje povprečno letno donosnost delnic podjetij obeh skupin. Trend donosnosti je zelo podoben za obe skupini, podobno kot pri finančnih kazalnikih pa skupina največjih 50 dosega boljše rezultate z izjemo leta 2021. S pomočjo standardnega odklona in variance donosnosti lahko ocenimo volatilnost donosnosti delnice skozi analizirano obdobje. Obe vrednosti sta izračunani na podlagi dnevni nihanj cen delnic v pripadajočem finančnem letu. Skupina največjih 50 dosega nižje vrednosti, kar je razvidno na sliki 12 in sliki 13, to pa nakazuje na nižje nihanje donosnosti.

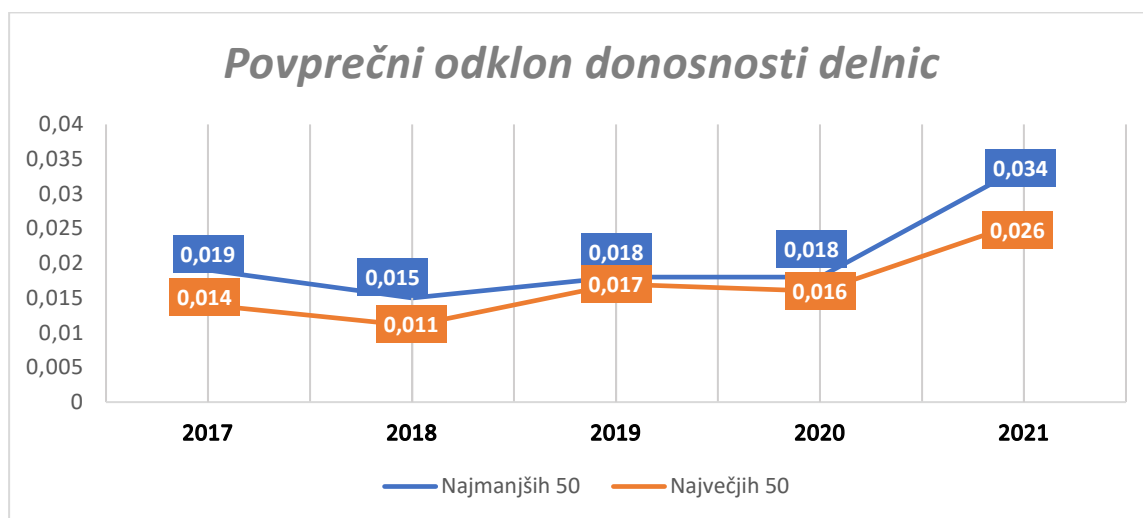
Primerjava tržnih kazalnikov pokaže podobne rezultate, kot jih je bilo moč opaziti pri finančnih kazalnikih in sicer da skupina največjih 50 podjetij dosega boljše povprečne rezultate. Tudi pri tržnih kazalnikih opazim, da je trend skozi 5 let zelo podoben za obe skupini.

Slika 11: Primerjava povprečne letne donosnosti delnic največjih 50 in najmanjših 50 podjetij



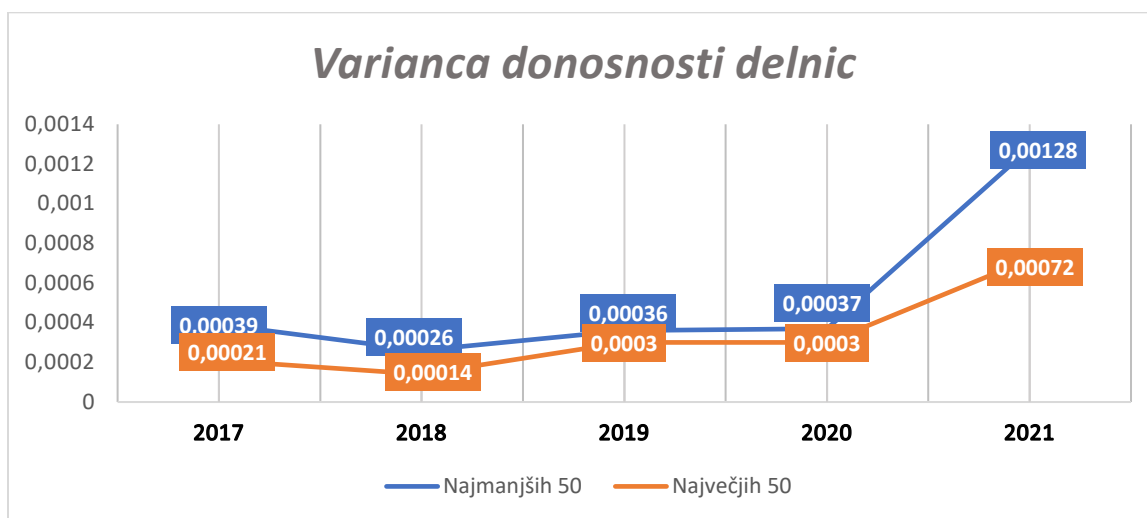
Vir: lastno delo, izračunano na podlagi tržnih podatkov podjetij pridobljenih preko Yahoo ! Finance API-ja.

Slika 12: Primerjava povprečnega standardnega odklona donosnosti delnic največjih 50 in najmanjših 50 podjetij



Vir: lastno delo, izračunano na podlagi tržnih podatkov podjetij pridobljenih preko Yahoo ! Finance API-ja.

Slika 13: Primerjava variance donosnosti delnic največjih 50 in najmanjših 50 podjetij

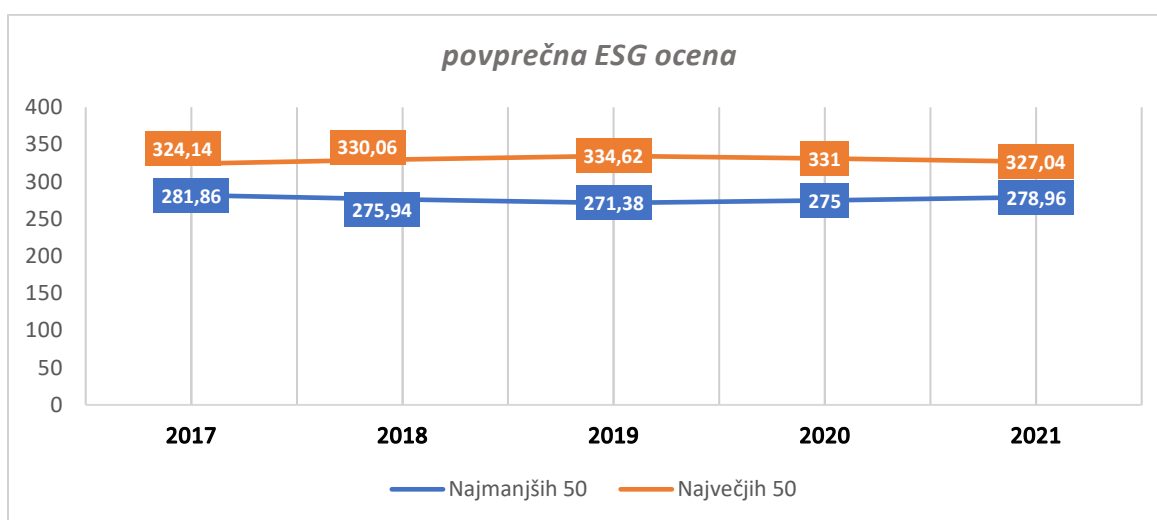


Vir: lastno delo, izračunano na podlagi tržnih podatkov podjetij pridobljenih preko Yahoo ! Finance API-ja.

#### 4.1.3 Primerjava rezultatov analize teksta

Spodnje slike 14-20 prikazujejo rezultate analize teksta po 6 izbranih merah, ki so predstavljene v poglavju 3.4. Glavni cilj te analize je bilo ugotoviti, ali dosega skupina največjih 50 podjetij tudi pri tej primerjavi boljše rezultate kot skupina najmanjših 50.

Slika 14: Primerjava povprečne ESG ocene največjih 50 in najmanjših 50 podjetij



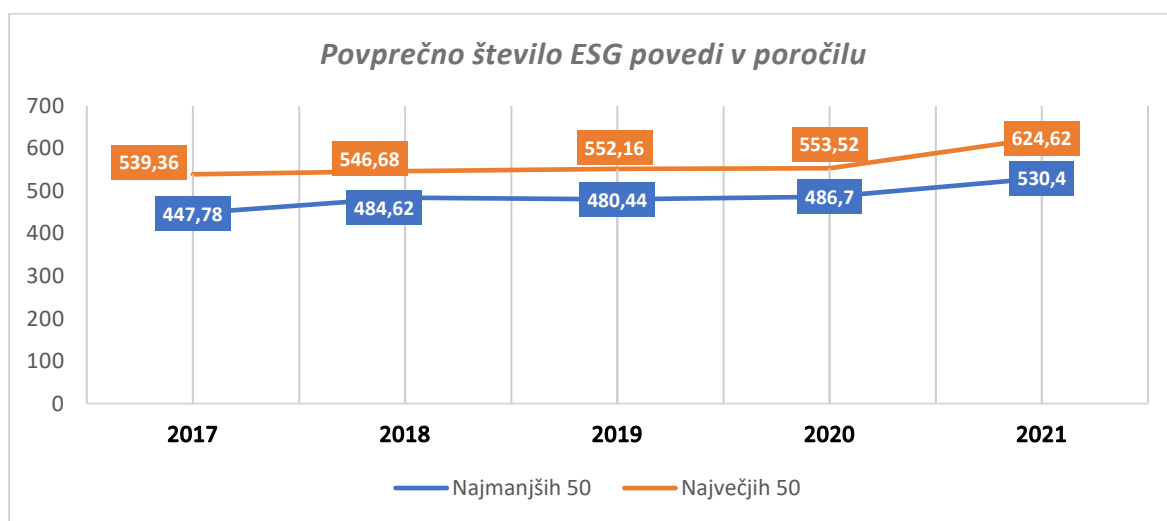
Vir: lastno delo.

Slika 14 primerja povprečno ESG oceno obeh raziskovanih skupin. Vidimo lahko, da skupina največjih 50 podjetij dosega boljše rezultate v vseh petih letih. Nihanje ocene skozi leta je majhno, saj je ocena odvisna le od povprečne uvrstitve podjetij na lestvici po merah

analize teksta. To pomeni, da ESG ocena ne odraža morebitnega padca ali rasti določenih ESG mer, če je pojav prisoten v celotnem analiziranem vzorcu.

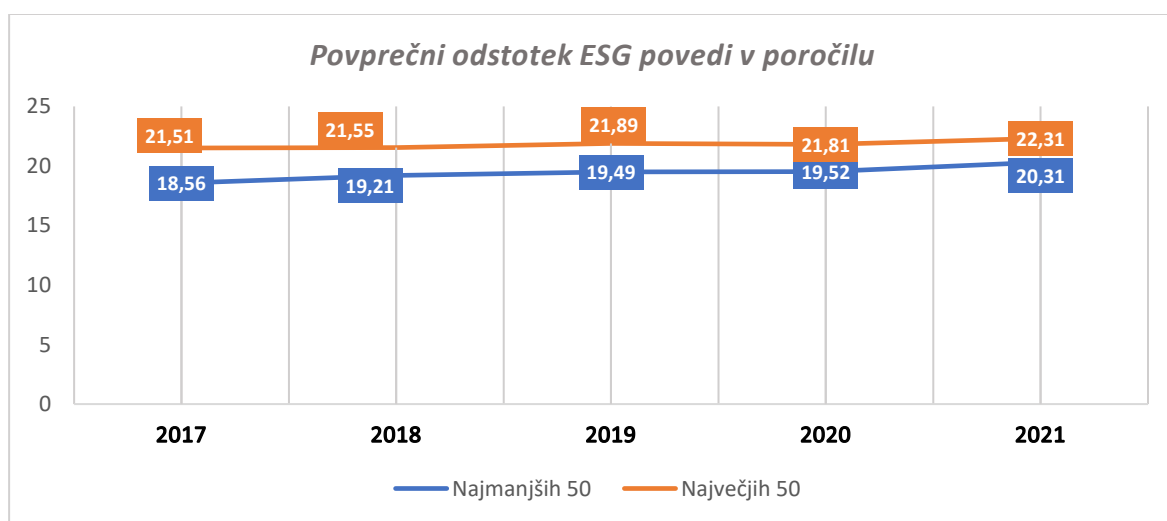
Sliki 15 in 16 prikazujeta primerjavo števila in odstotka ESG povedi v poročilih. Vidno je, da sta ti vrednosti višji pri skupini največjih 50 in to za vseh 5 let. Opazimo lahko še, da vrednosti nekoliko naraščata skozi leta, kar pomeni, da podjetja z leti posvečajo večji del poročila ESG temam.

*Slika 15: Primerjava povprečnega števila ESG povedi v poročilu največjih 50 in najmanjših 50 podjetij*



*Vir: lastno delo.*

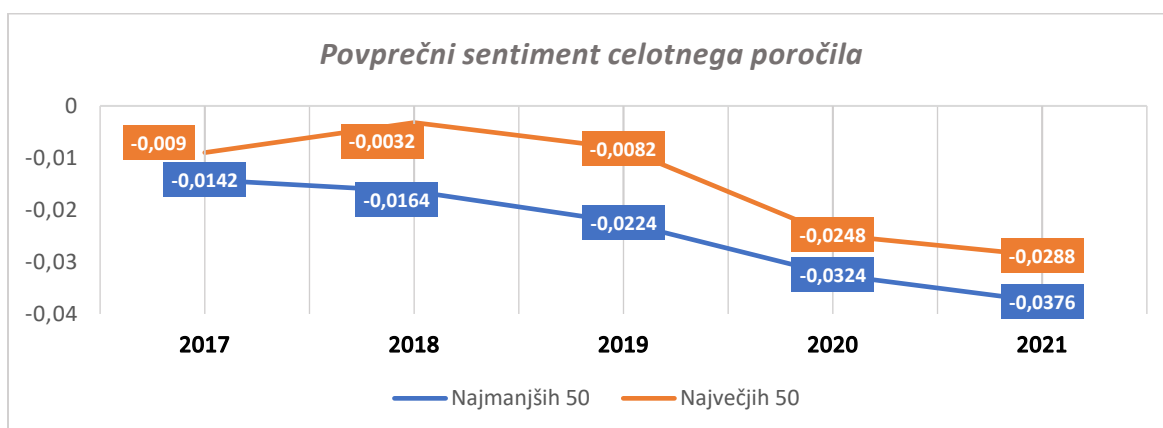
*Slika 16: Primerjava povprečnega odstotka ESG povedi v poročilu največjih 50 in najmanjših 50 podjetij*



*Vir: lastno delo.*

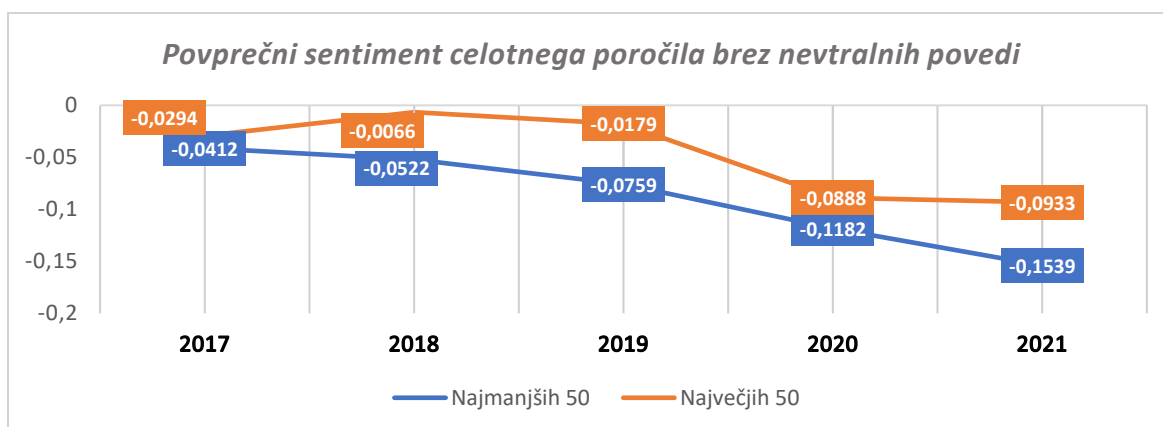
Slika 17 prikazuje povprečni sentiment ali povprečno razpoloženje v povedih celotnega poročila. Model za napovedovanje razpoloženja vrne vrednost 1 za povedi, kjer je moč zaznati pozitivno razpoloženje, 0 za povedi, kjer ni zaznano ne pozitivno in ne negativno razpoloženje torej nevtralne povedi ter -1 za povedi, kjer prevladuje negativno razpoloženje. Povprečni sentiment lahko tako zavzame vrednosti na intervalu:  $[-1, 1]$ . Slika 18 prikazuje povprečni sentiment poročila, če izločimo povedi, kjer ne zaznamo pozitivnega ali negativnega razpoloženja. Takšna izločitev omogoča boljše razumevanje razmerja med pozitivnimi in negativnimi povedmi. Pri obeh vrednostih povprečnega sentimenta podjetja iz skupine največjih 50 dosegajo boljši rezultat za vseh pet let, torej so letna poročila teh podjetij nekoliko bolj pozitivno naravnana oziroma glede na negativne vrednosti manj negativno naravnana. Zanimivo je opaziti podoben trend pri obeh skupinah, saj vrednosti od leta 2018 naprej padajo, kar pomeni, da z leti pada zaznano splošno razpoloženje v poročilih.

*Slika 17: Primerjava povprečnega sentimenta celotnega poročila največjih 50 in najmanjših 50 podjetij*



*Vir: lastno delo.*

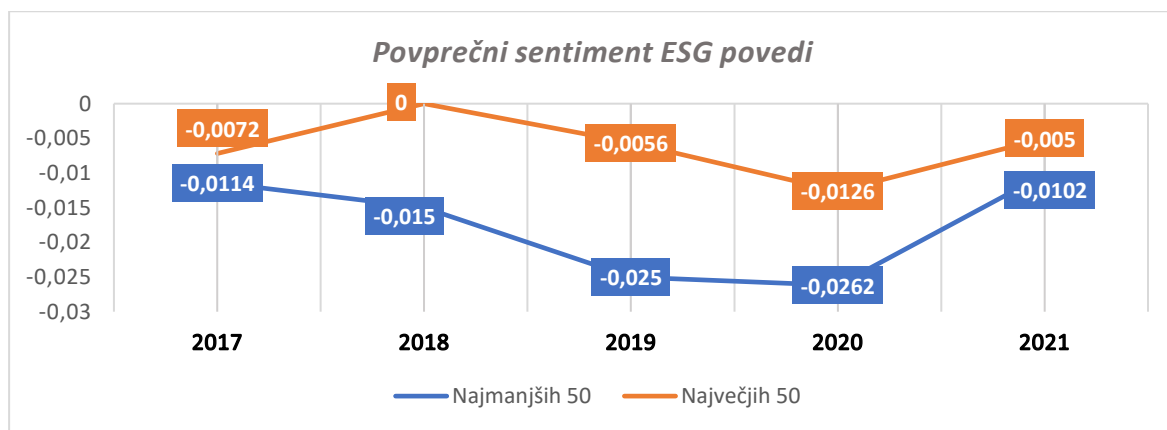
*Slika 18: Primerjava povprečnega sentimenta celotnega poročila brez nevtralnih povedi največjih 50 in najmanjših 50 podjetij*



*Vir: lastno delo.*

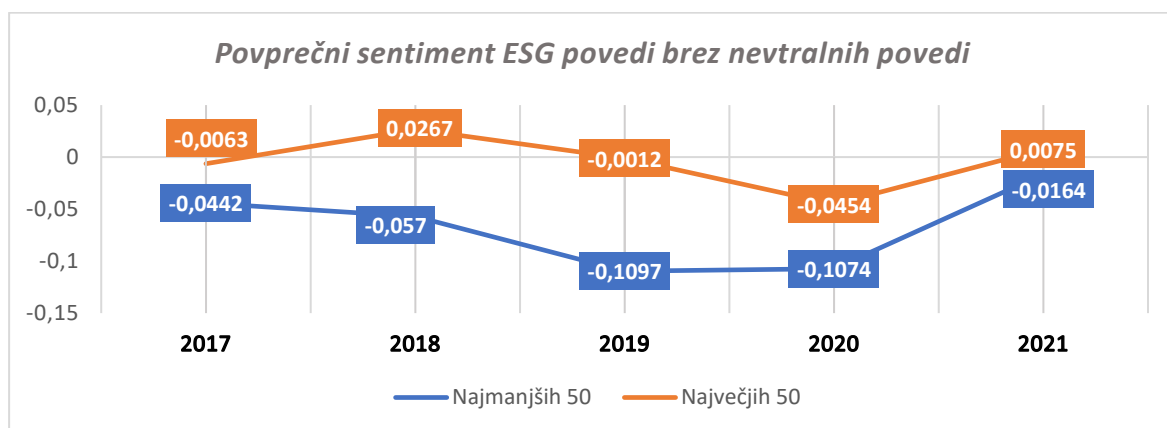
Naslednji analizirani meri se osredotočata na razpoloženje v povedih, ki se navezujejo na dejavnike ESG. Vrednosti za obe skupini so prikazane na slikah 19 in 20. Podobno kot pri razpoloženju celotnega besedila, tudi ko se omejimo samo na ESG povedi, skupina največjih 50 dosega višje vrednosti, ki nakazujejo na bolj pozitivno oziroma manj negativno razpoloženje. Tudi trend je ponovno zelo podoben za obe skupini. Dodatno je zanimivo, da je trend zelo podoben kot pri nekaterih finančnih kazalnikih (rast prihodkov, bruto marža, neto marža in marža iz poslovanja), kjer smo lahko opazili padec od leta 2018 do 2020 in nato rast v letu 2021. To povezavo oziroma podobnost v trendu v nadaljevanju magistrskega dela dodatno preverim z izračunom korelacije med posameznimi merami ESG ocene in finančnimi kazalniki.

Slika 19: Primerjava povprečnega sentimenta ESG povedi največjih 50 in najmanjših 50 podjetij



Vir: lastno delo.

Slika 20: Primerjava povprečnega sentimenta ESG povedi brez nevtralnih povedi največjih 50 in najmanjših 50 podjetij



Vir: lastno delo.

Ključna ugotovitev po primerjavi rezultatov analize teksta je ta, da skupina največjih 50 dosega višje povprečne vrednosti pri prav vseh ESG merah za vsa analizirana leta. Nadalje



opažena podobnosti trenda govori o tem, da so podjetja obeh skupin v enaki oziroma zelo podobni meri podvržena vplivom trga. Največjo razliko med skupinama je moč opaziti pri kategoriji ESG sentiment.

Še posebej zanimiva je opažena razlika v trendu med sentimentom celotnega poročila in ESG sentimentom. Pri obeh skupinah je splošno razpoloženje v poročilih od leta 2018 naprej padalo, medtem ko se razpoloženje prisotno v ESG povedih od leta 2020 do 2021 dvigne oziroma postane bolj pozitivno. To razlikovanje je spodbudilo raziskovanje povezave med razpoloženjem v ESG delih poročil in delih, ki se ne navezujejo na dejavnike ESG. Ker so ESG povedi podmnožica vseh povedi poročila, sem pričakoval, da bo vrednost ESG razpoloženja le malo drugačna od skupnega razpoloženja. Za boljše razumevanje te povezave izračunam koeficient korelacije med razpoloženjem v povedih, ki se navezujejo na ESG teme in povedih, ki se ne. Povprečno korelacijo izračunam po metodi predstavljeni v poglavju 3.6. Korelacijski koeficienti so izračunani z uporabo Python knjižnice SciPy. Nizko povezavo bi lahko razumeli kot pomembno razlikovanje v razpoloženju med ESG in NE-ESG deli poročila. Rezultat je prikazan v tabeli 5.

*Tabela 5: Povprečna korelacija med razpoloženjem v povedih, ki se navezujejo na ESG in tistih, ki se ne*

	Sentiment NE ESG povedi	Sentiment ESG povedi
Sentiment NE ESG povedi	1	0,526
Sentiment ESG povedi	0,526	1

*Vir: lastno delo.*

Korelacijski koeficient med razpoloženjem v povedih, ki se navezujejo na ESG teme in tistih, ki se ne, je 0,526, kar nakazuje le na zmerno povezanost. Na razliko pri razpoloženju med skupnim sentimentom in ESG sentimentom bom še posebej pozoren pri analizi ESG ocene po panogah, kjer raziskujem povezavo med sentimentom in ESG tveganjem.

#### **4.2 Primerjava najboljših 50 in najslabših 50 podjetij po ESG oceni**

V drugem sklopu primerjav se odločim za drugačno razporeditev podjetij. Namesto povprečne tržne kapitalizacije podjetja razdelim glede na povprečno ESG oceno skozi analiziranih 5 let. Zanimalo me je dvoje, ali bo v skupini 50 najboljših velika večina podjetij iz skupine 50 največjih in ali bo takšna delitev ohranila razlike pri finančnih in tržnih kazalnikih. Spodnja tabela 6 prikazuje podjetja razdeljena glede na povprečno ESG oceno.

Tabela 6: Skupini najboljših 50 in najslabših 50 podjetij glede na povprečno ESG oceno

Najboljših 50 podjetij glede na ESG oceno			Najslabših 50 podjetij glede na ESG oceno		
Podjetje	ESG ocena	Skupina glede na tržno kapitalizacijo	Podjetje	ESG ocena	Skupina glede na tržno kapitalizacijo
Intel Corporation	546	največjih 50	Sealed Air Corporation	302	najmanjših 50
Chevron Corporation	525	največjih 50	Philip Morris International	301	največjih 50
Henry Schein	504	najmanjših 50	Alphabet	301	največjih 50
Merck & Co	501	največjih 50	Juniper Networks	300	najmanjših 50
Verizon Communications	497	največjih 50	Pfizer	298	največjih 50
Adobe	487	največjih 50	Honeywell	297	največjih 50
FMC	487	najmanjših 50	Amgen	294	največjih 50
Visa	487	največjih 50	International Flavors & Fragrances	283	najmanjših 50
AT&T	482	največjih 50	IBM	281	največjih 50
3M	478	največjih 50	Howmet Aerospace	281	najmanjših 50
Eastman Chemical Company	465	najmanjših 50	Tractor Supply Company	272	najmanjših 50
Home Depot	462	največjih 50	Robert Half International	255	najmanjših 50
Medtronic	461	največjih 50	Marathon Oil	250	najmanjših 50
Exxon Mobil Corporation	451	največjih 50	DaVita	244	najmanjših 50
Procter & Gamble Company	440	največjih 50	Masco	243	najmanjših 50
Bristol-Myers	440	največjih 50	PerkinElmer	236	najmanjših 50
C. H. Robinson Worldwide	437	najmanjših 50	News Corporation	232	najmanjših 50
F5	436	najmanjših 50	Coterra Energy	226	najmanjših 50
Microsoft	436	največjih 50	Iron Mountain	224	najmanjših 50
Zions Bancorporation	435	najmanjših 50	Globe Life	223	najmanjših 50
Eli Lilly and Company	414	največjih 50	Boeing	220	največjih 50
Nielsen	411	najmanjših 50	PayPal	219	največjih 50
McDonald's Corporation	404	največjih 50	Ralph Lauren	213	najmanjših 50
Comcast Corporation	395	največjih 50	CF Industries Holdings	204	najmanjših 50
Johnson & Johnson	393	največjih 50	Walt Disney Company	200	največjih 50
Wells Fargo	392	največjih 50	Assurant	190	najmanjših 50
PVH Corp	388	najmanjših 50	Costco Wholesale Corporation	184	največjih 50
Citigroup	381	največjih 50	PulteGroup	183	najmanjših 50
Salesforce.com	380	največjih 50	Mosaic	181	najmanjših 50
AES Corp	378	najmanjših 50	Quanta Services	179	najmanjših 50
PepsiCo	377	največjih 50	Broadcom	175	največjih 50
Mastercard Incorporated	375	največjih 50	Dentsply Sirona	175	najmanjših 50
Snap-on Incorporated	369	najmanjših 50	Bath & Body Works	171	najmanjših 50
United Health Group Incorporated	359	največjih 50	Kimco Realty	169	najmanjših 50
Borg Warner	347	najmanjših 50	Bank of America	168	največjih 50
J.B. Hunt Transport Services	345	najmanjših 50	Jacobs	166	najmanjših 50
Interpublic Group	340	najmanjših 50	Pentair	166	najmanjših 50

se nadaljuje

Tabela 6: Skupini najboljših 50 in najslabših 50 podjetij glede na povprečno ESG oceno (nad.)

Najboljših 50 podjetij glede na ESG oceno			Najslabših 50 podjetij glede na ESG oceno		
Podjetje	ESG ocena	Skupina glede na tržno kapitalizacijo	Podjetje	ESG ocena	Skupina glede na tržno kapitalizacijo
Berkshire Hathaway	340	največjih 50	Under Armour	165	najmanjših 50
Walmart	338	največjih 50	Meta	162	največjih 50
Hasbro	334	najmanjših 50	NRG Energy	161	najmanjših 50
Coca-Cola Company	334	največjih 50	Amazon	160	največjih 50
Qorvo	328	najmanjših 50	Advance Auto Parts	143	najmanjših 50
Allegion	327	najmanjših 50	Textron	143	najmanjših 50
Oracle	322	največjih 50	Accenture	136	največjih 50
United Rentals	318	najmanjših 50	Netflix	132	največjih 50
AbbVie	316	največjih 50	Nike	126	največjih 50
Nvidia	316	največjih 50	Avery Dennison	104	najmanjših 50
Abbott Laboratories	315	največjih 50	Comerica Incorporated	100	najmanjših 50
Invesco	305	najmanjših 50	JPMorgan	89	največjih 50
Universal Health Services	303	najmanjših 50	Apple	72	največjih 50

Vir: lastno delo, izračunano na podlagi podatkov pridobljenih iz <https://companiesmarketcap.com>.

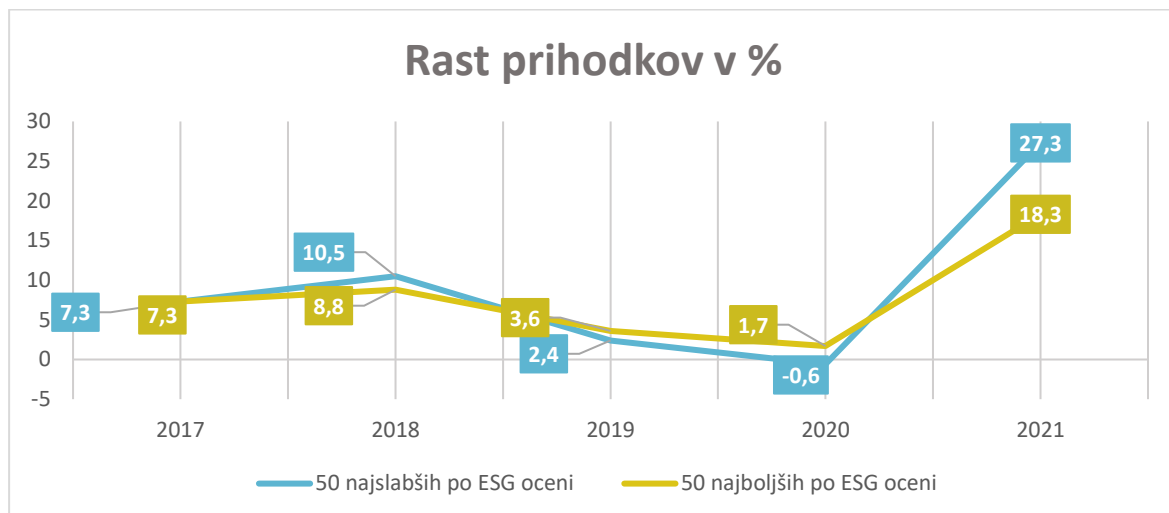
Lahko opazimo, da je v skupini najboljših 50 po ESG oceni 31 podjetij iz skupine največjih 50 in 19 iz skupine najmanjših 50. V odstotkih je to 62% proti 38% v korist skupine največjih 50. Takšna delitev je pričakovana glede na to, da so bila povprečja mer analize teksta višja pri vseh primerjavah za največjih 50 podjetij. Vseeno pa nova delitev kar dobro premeša podjetja in naredi primerjavo finančnih ter tržnih kazalnikov toliko bolj zanimivo.

#### 4.2.1 Primerjava finančnih kazalnikov

Spodnje slike 21-27 prikazujejo primerjavo finančnih in tržnih kazalnikov v enakem zaporedju kot v poglavju 4.1.1. Slika 21 primerja skupini po kazalniku rast prihodkov. Vidimo, da so vrednosti za obe skupini precej podobne z izjemo zadnjega leta, kjer je skupina najslabših 50 po ESG oceni boljša za 9%.

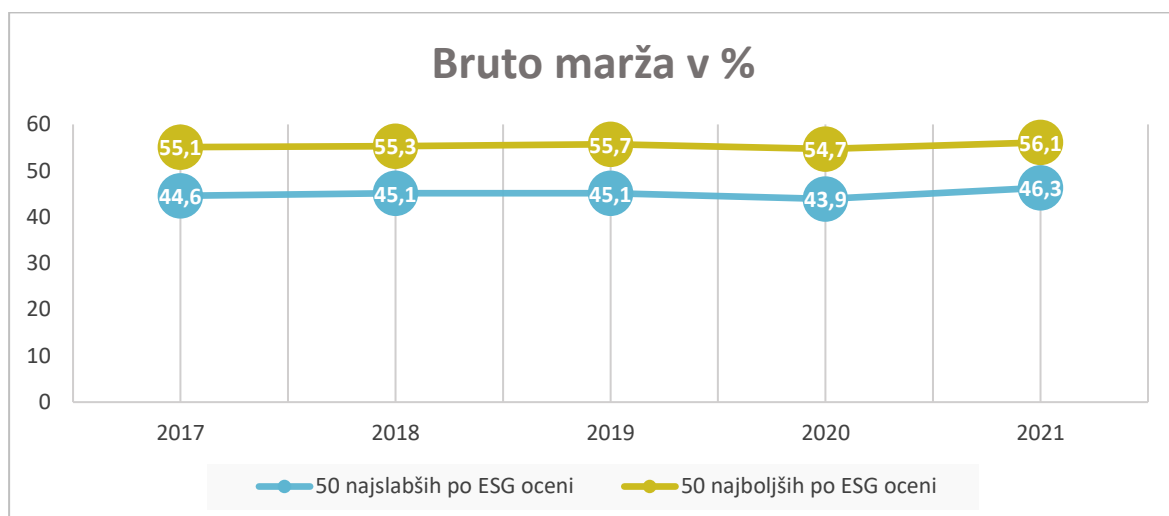
Slike 22, 23 in 24 prikazujejo primerjavo kazalnikov bruto marža, neto marža in marža prostega denarnega toka. Pri vseh treh kazalnikih je skupina 50 najboljših podjetij glede na ESG oceno v prednosti. Trendi obeh skupin so zelo podobni. Razlika med skupinama je nekoliko manjša kot pri prvotni delitvi glede na povprečno tržno kapitalizacijo.

Slika 21: Primerjava rasti prihodkov najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

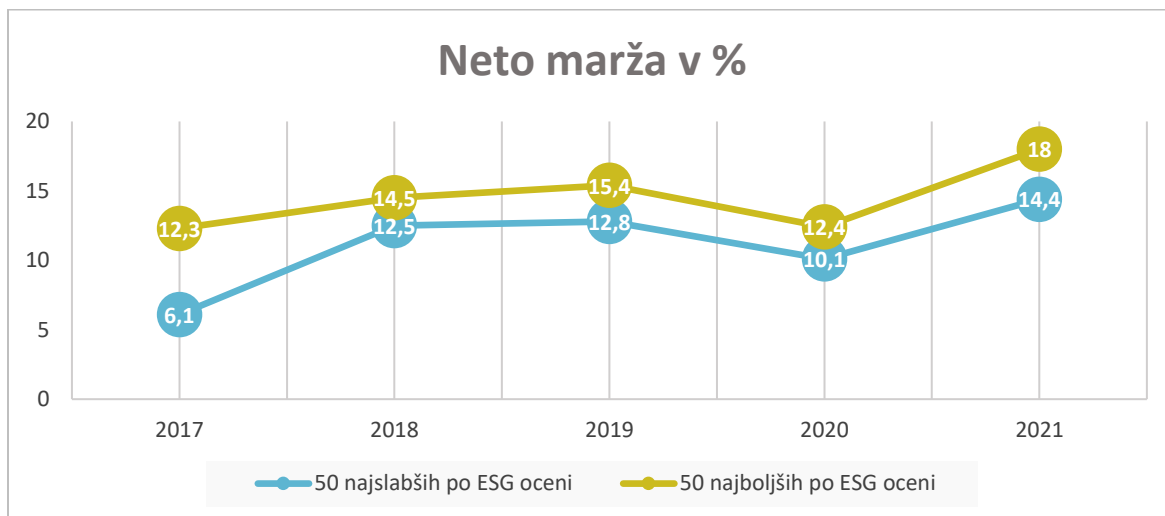
Slika 22: Primerjava bruto marže najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

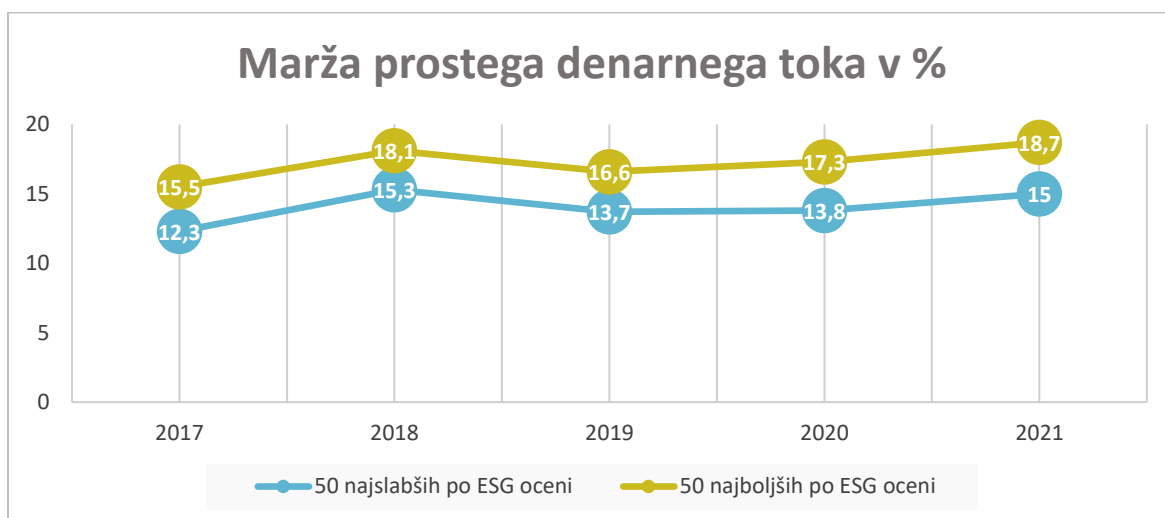
Tudi primerjava marže iz poslovanja daje prednost skupini 50 najuspešnejših po ESG oceni. To je vidno na sliki 25.

Slika 23: Primerjava neto marže najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

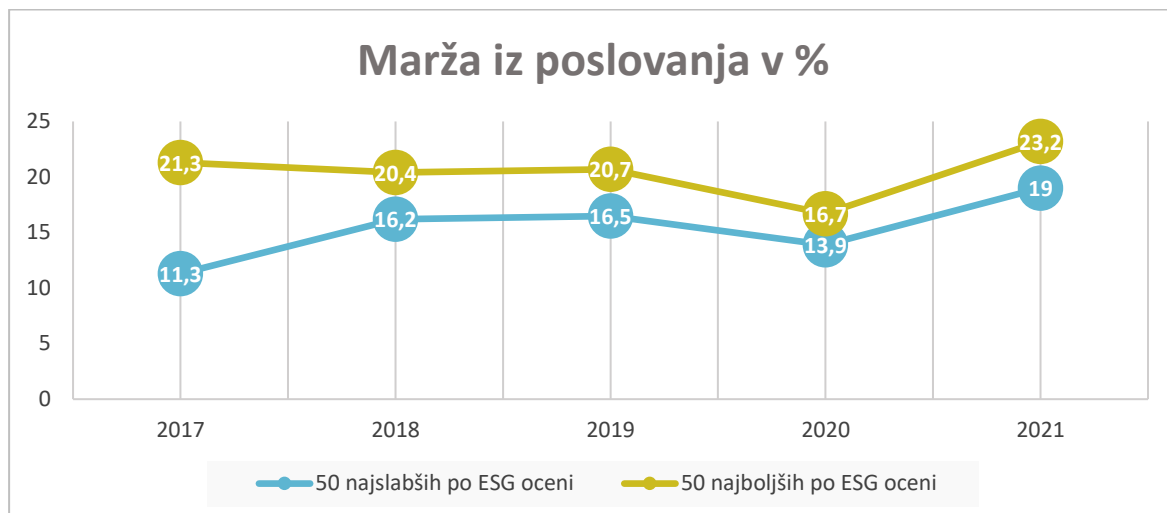
Slika 24: Primerjava marže prostega denarnega toka najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

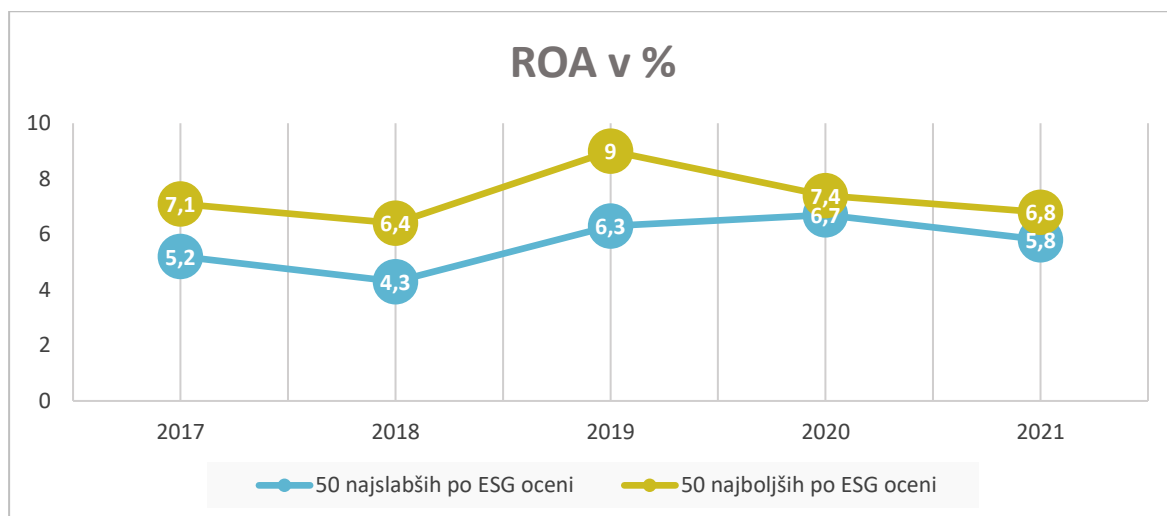
Sliki 26 in 27 prikazujeta primerjavo kazalnikov ROA in ROE. Pri ROA je skupina najboljših po ESG v prednosti v vseh petih letih, medtem ko pri ROE skupina najslabših dosega boljše rezultate v letih 2020 in 2021.

Slika 25: Primerjava marže iz poslovanja najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

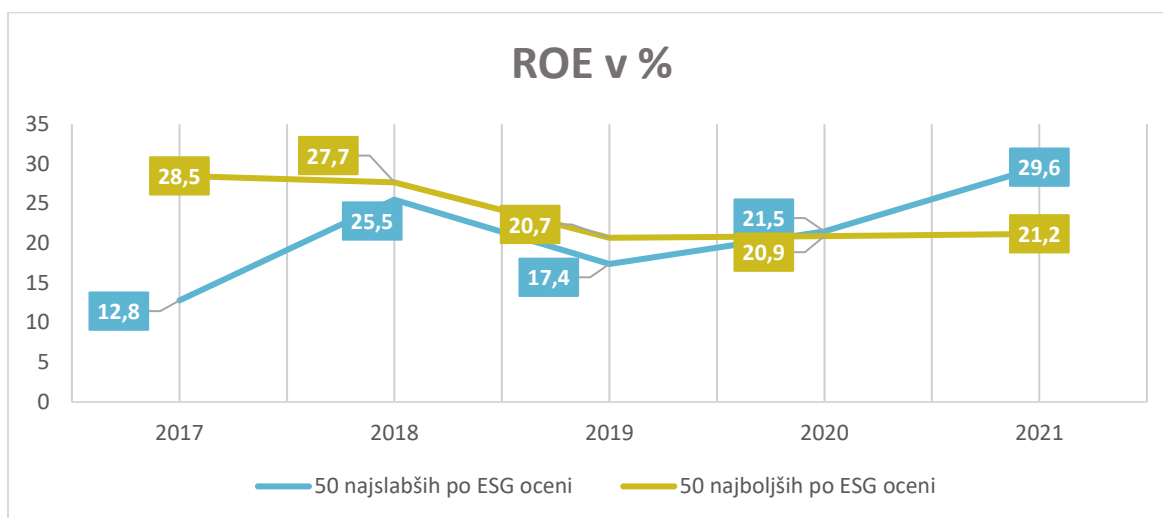
Slika 26: Primerjava ROA najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

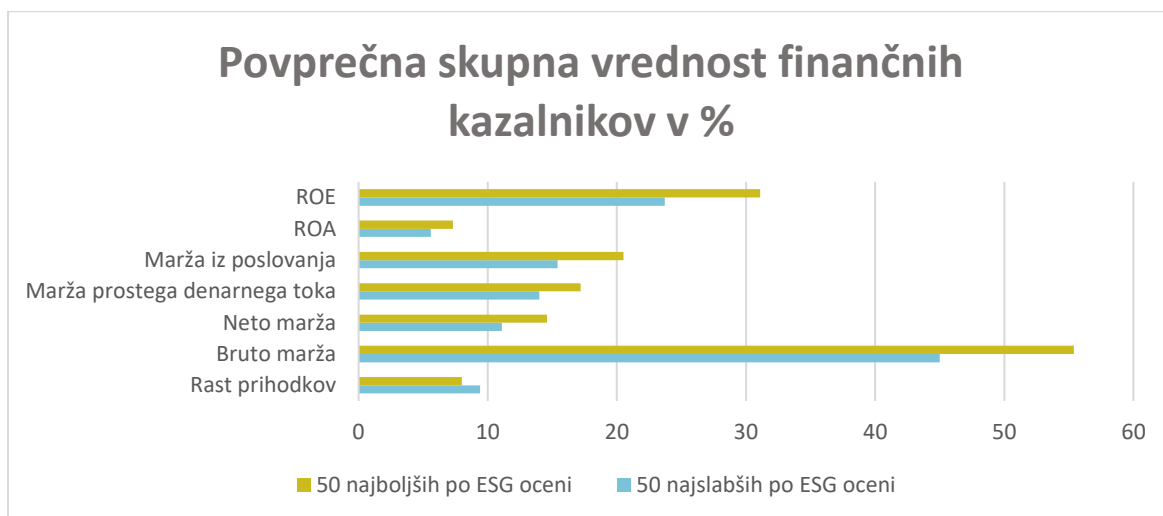
Na sliki 28 je razvidno, da skupina po ESG oceni najboljših dosega boljše rezultate pri vseh finančnih kazalnikih z izjemo rasti prihodkov. Razlike so nekoliko manjše kot pri primerjavi največjih in najmanjših po tržni kapitalizaciji. Kljub temu pa ostaja razlika med skupinama očitna.

Slika 27: Primerjava ROE najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

Slika 28: Primerjava povprečne skupne vrednosti finančnih kazalnikov najboljših 50 in najslabših 50 podjetij glede na ESG oceno

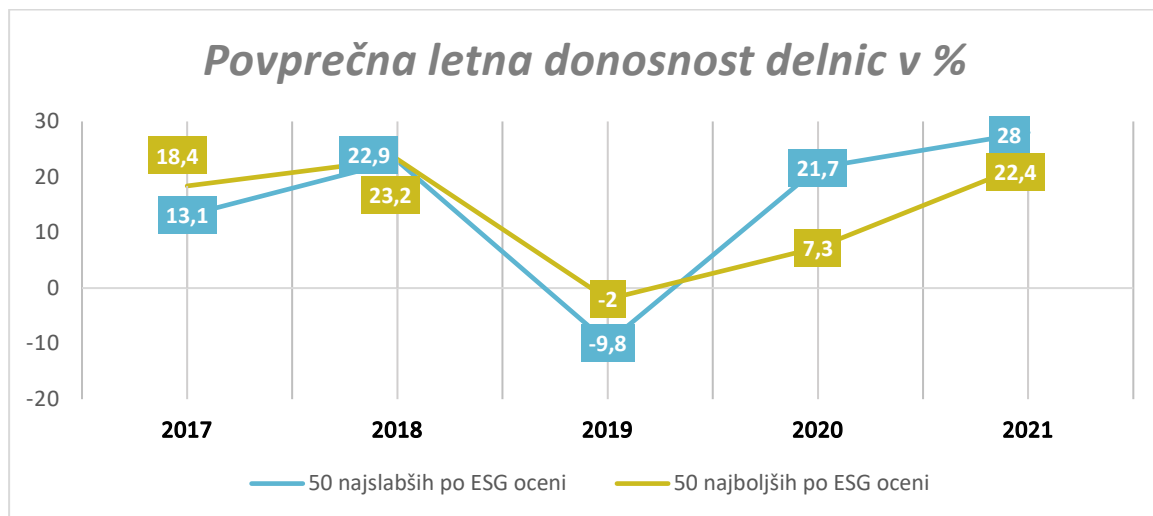


Vir: lastno delo, izračunano na podlagi finančnih izkazov podjetij pridobljenih iz <https://stockanalysis.com/>.

#### 4.2.2 Primerjava tržnih kazalnikov

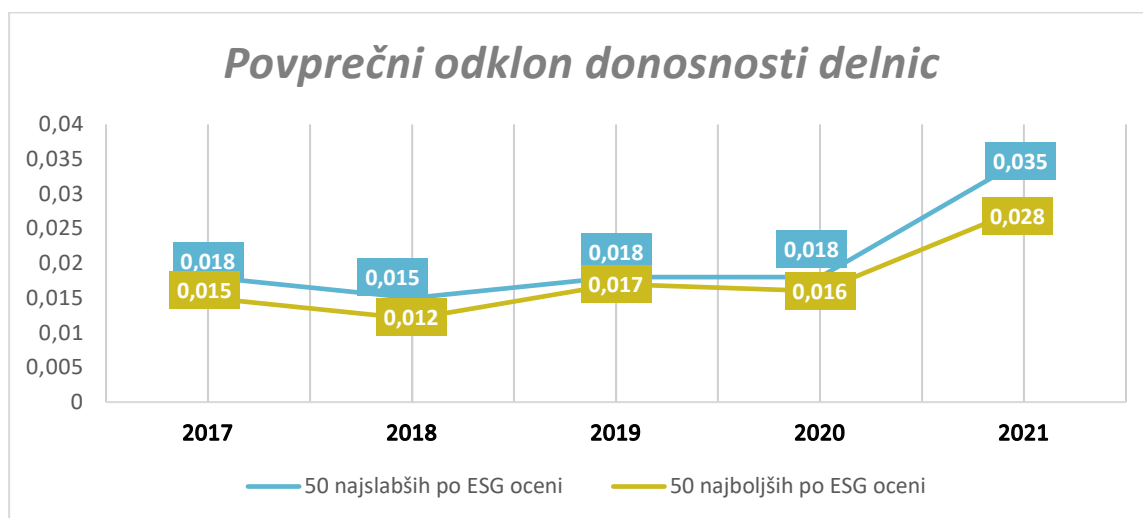
Slike 29, 30 in 31 prikazuje primerjavo kazalnikov izračunanih na podlagi tržnih podatkov. Opazimo lahko, da se skupini izmenjujeta v prednosti pri kazalniku povprečna donosnost delnic. Je pa pri skupini 50 ESG uspešnejših moč opaziti nižje nihanje donosnosti oziroma nižjo volatilitnost.

Slika 29: Primerjava povprečne letne donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi tržnih podatkov pridobljenih preko Yahoo ! Finance API-ja.

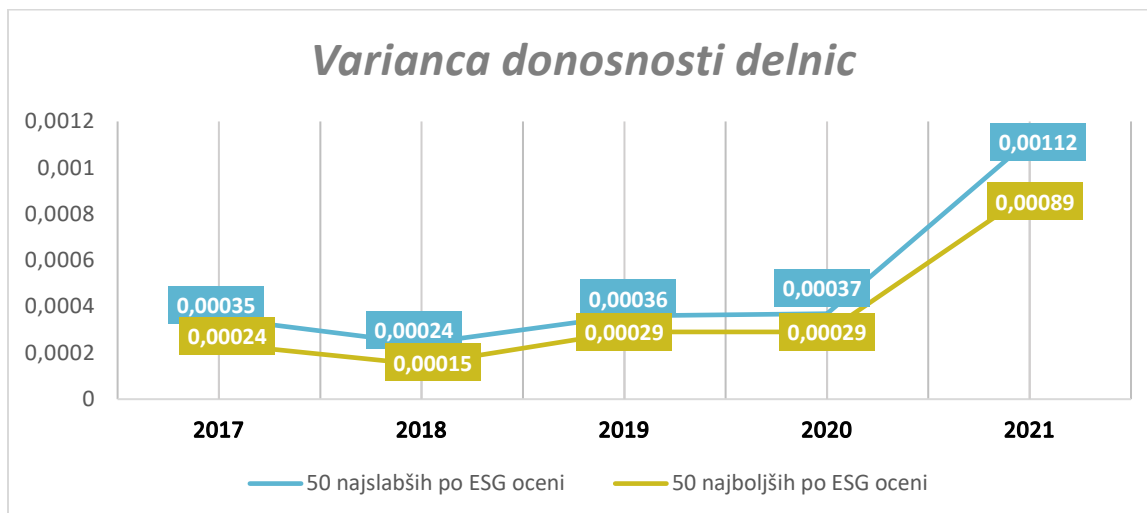
Slika 30: Primerjava povprečnega odklona donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi tržnih podatkov pridobljenih preko Yahoo ! Finance API-ja.



Slika 31: Primerjava variance donosnosti delnic najboljših 50 in najslabših 50 podjetij glede na ESG oceno



Vir: lastno delo, izračunano na podlagi tržnih podatkov pridobljenih preko Yahoo ! Finance API-ja.

#### 4.3 Analiza korelacije med ESG oceno in finančno uspešnostjo podjetij

Podjetja iz skupine največjih 50 po tržni kapitalizaciji dosegajo višje ESG ocene. Nadaljnja delitev podjetij glede na ESG oceno pokaže, da imajo po ESG oceni bolj uspešna podjetja boljše finančne rezultate. Zaradi teh dveh ugotovitev me je zanimalo, ali lahko s statističnimi metodami zaključim, da obstaja povezava med ESG oceno podjetij in finančno uspešnostjo.

Pri izračunu povprečne vrednosti korelacijskih koeficientov ponovno sledim metodologiji predstavljeni v poglavju 3.6. S tem izračunom želim preveriti, kakšna je korelacija med rezultati, ki jih v določenem letu posamezno podjetje dosega pri merah analize teksta in finančno uspešnostjo izraženo z definiranimi kazalniki. Rezultati so predstavljeni v spodnji tabeli 7.

Tabela 7: Matrika povprečnih korelacijskih koeficientov med ESG oceno, posameznimi deli ESG ocene in kazalniki finančne uspešnosti

	Rast prihodkov	Bruto marža	Neto marža	Marža prostega denarnega toka	ROA	ROE	Marža iz poslovanja	Povprečna donosnost delnic	Odklon donosnosti delnic	Variance donosnosti delnic
ESG ocena	-0,04	0,04	0,06	0,07	0,06	0,14	0,02	0,12	-0,05	-0,05
Število ESG povedi	0,23	0,1	0,13	0,00	-0,12	0,01	0,1	0,04	0,07	0,06

se nadaljuje

Tabela 7: Matrika povprečnih korelacijskih koeficientov med ESG oceno, posameznimi deli ESG ocene in kazalniki finančne uspešnosti (nad.)

	Rast prihodkov	Bruto marža	Neto marža	Marža prostega denarnega toka	ROA	ROE	Marža iz poslovanja	Povprečna donosnost delnic	Odklon donosnosti delnic	Varianca donosnosti delnic
<b>Odstotek ESG povedi</b>	0,06	0,23	0,25	0,13	-0,11	0,07	0,17	0,04	0,35	0,35
<b>Sentiment celotnega poročila</b>	0,01	0,01	0,08	-0,13	0,06	0,09	-0,04	-0,03	-0,39	-0,39
<b>Sentiment celotnega poročila brez nevtralnih</b>	-0,03	0,03	-0,09	-0,1	0,05	0,05	-0,06	-0,05	-0,44	-0,44
<b>ESG sentiment</b>	0,15	0,09	0,12	0,05	-0,04	-0,01	0,11	0,1	0,14	0,14
<b>ESG sentiment brez nevtralnih</b>	0,14	0,17	0,18	0,07	-0,06	0,03	0,13	0,18	0,1	0,12

Vir: lastno delo.

Pri korelacijah med kazalniki finančne uspešnosti in ESG oceno lahko opazimo, da so povprečne vrednosti koeficientov zelo nizke. Nekoliko višje vrednosti lahko opazimo pri posameznih merah ESG ocene. Pri številu ESG povedi v poročilu izstopa povezava z rastjo prihodkov, a je povprečna vrednost koeficienta (0,23) še vedno nizka. Največ višjih vrednosti lahko vidimo pri meri odstotek ESG povedi, kjer opažam šibko pozitivno korelacijo z bruto maržo (0,23) in neto maržo (0,25) ter zmerno pozitivno korelacijo z odklonom (0,35) in varianco (0,35) donosnosti delnic. Pozitivna korelacija pri odklonu in varianci donosnosti nakazuje na višjo volatilitnost, torej da je višji odstotek ESG povedi v poročilu povezan z večjim nihanjem donosnosti delnic. Obratno je moč opaziti pri merah povprečnega sentimenta v poročilu, kjer je korelacija z odklonom in varianco donosnosti še močnejša in negativna, torej da je višje zaznano razpoloženje v poročilu povezano z manjšim nihanjem donosnosti. Pri merah ESG sentimenta lahko opazimo, da je več vrednosti višjih od 0,15 (rast prihodkov, bruto marža, neto marža in povprečna donosnost delnic) in da so vse te vrednosti pozitivne, torej da obstaja šibka povezava med višjim ESG razpoloženjem in višjo finančno uspešnostjo. Če se omejim samo na višje negativne vrednosti, ki bi nakazovale, da obstaja povezava med višjimi vrednostmi ESG mer in nižjo finančno uspešnostjo, jih lahko zaznam samo pri kazalniku ROA (s številom in odstotkom ESG povedi) in marži prostega denarnega toka (z merama sentimenta celotnega poročila).

Nekatere sorodne študije ugotavljajo, da obstaja pozitivna povezava med ESG in CFP z več letnim zamikom. Torej da se višja ESG uspešnost odraža v višji finančni uspešnosti čez

nekaj let. Shinohara (2014) postavi to mejo na 5 let. Sam se v svoji raziskavi omejim na 5 finančnih let, zato preverba te ugotovitve ni mogoča. Vsak zamik enega let znotraj pridobljenih podatkov zmanjša nabor vrednosti za 100 letnih poročil in finančnih izkazov. Iz tega razloga se za ta test ne odločim.

#### 4.4 Primerjava med gospodarskimi panogami

Izbira 100 podjetij iz skupine S&P 500 omogoča tudi delitev glede na GICS panogo (Globalni standard za klasifikacijo panoge). Standard določa 11 panog vidnih v tabeli 8:

*Tabela 8: GICS panoge*

<b>GICS panoge</b>		
Energetika	Zdravstvo	Informacijska tehnologija
Nujna potrošnja	Komunikacijske storitve	Industrija
Surovine	Javne storitve	Finance
Ne nujna potrošnja	Nepremičnine	

*Prirejeno po Admirals (2022).*

Smiselno je torej preveriti, kako se med panogami razlikuje ESG ocena oziroma posamezne mere ESG ocene. Dodatno me je zanimalo, ali so ocene nižje za panoge oziroma podjetja, ki so izpostavljena večjim ESG tveganjem. Ker so bila podjetja izbrana na podlagi povprečne tržne kapitalizacije, so panoge neenakomerno zastopane. Spodnja tabela 9 prikazuje število pripadajočih podjetij za vsako od enajstih panog.

*Tabela 9: Število pripadajočih podjetij za vsako GICS panogo*

<b>Panoga</b>	<b>Število podjetij</b>
Energetika	4
Zdravstvo	15
Informacijska tehnologija	16
Nujna potrošnja	6
Komunikacijske storitve	9
Industrija	16
Surovine	7
Javne storitve	2
Finance	10
Ne nujna potrošnja	13
Nepremičnine	2

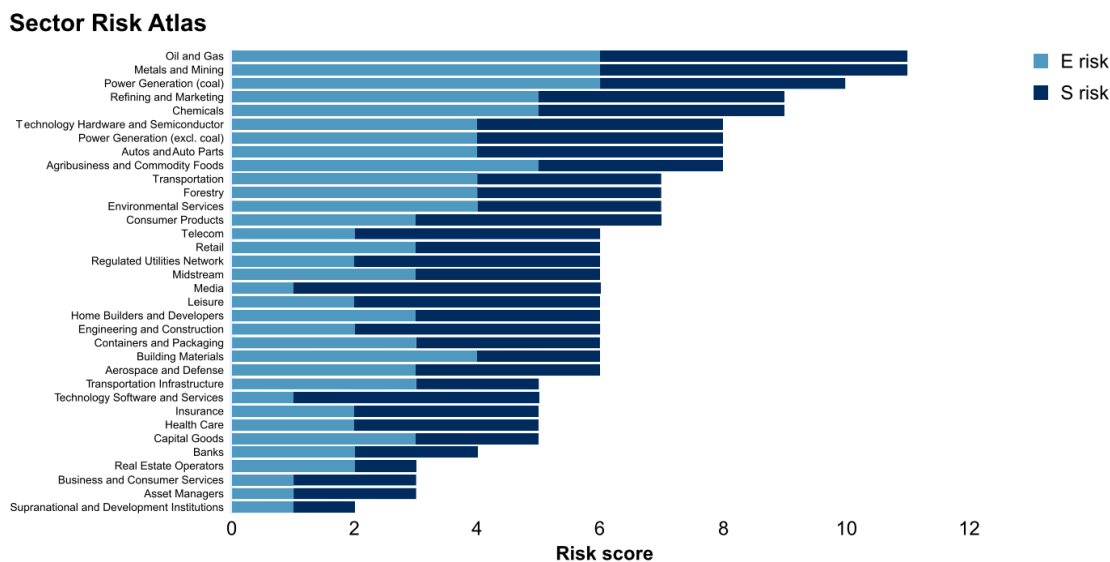
*Vir: lastno delo.*

Ocena panog z manj predstavniki (recimo Nepremičnine in Javne storitve z dvema predstavnikoma) bo toliko bolj odvisna od ocene posameznih pripadajočih podjetij. Pri

panogah z manj predstavnikov tako težko sprejmemo ESG oceno kot verodostojen pokazatelj ESG uspešnosti celotne panoge.

Slika 32 prikazuje oceno tveganja, ki jo posamezni dejavnosti dodelijo pri S&P Global.

*Slika 32: Razvrstitev gospodarskih dejavnosti po E in S tveganju*



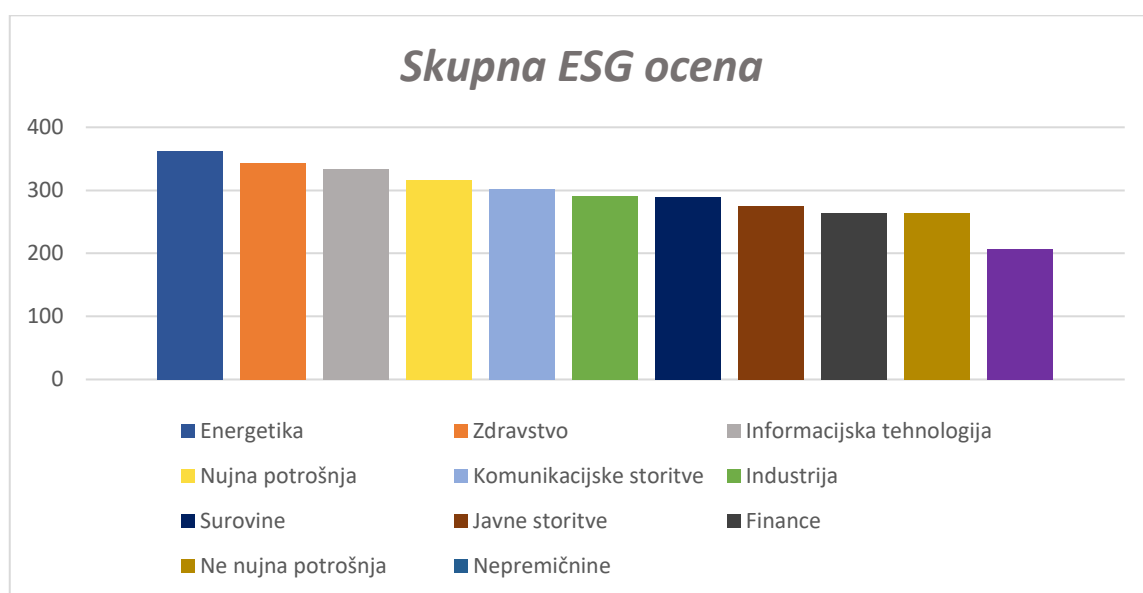
*Vir: S&P Global (2019).*

Ocene v zgornji tabeli uporabim kot referenco za oceno ESG tveganja dejavnosti in primerjavo z ESG ocenami pridobljenimi v moji raziskavi. V tabeli tveganja niso razvrščena po GICS panogah, a kljub temu lahko vidimo, da je višje tveganje prisotno v panogah Energetika, Industrija in Javne storitve ter nižje tveganje v panogah Informacijska tehnologija, Zdravstvo in Nepremičnine.

Slika 33 prikazuje povprečno skupno ESG oceno panogi pripadajočih podjetij, ki so bila vključena v analizo. Prva tri mesta zasedajo panoge Energetika, Zdravstvo in Informacijska tehnologija, medtem ko so pri repu Nepremičnine, Ne nujna potrošnja in Finance. V primerjavi z zgornjo sliko 32 je ta razvrstitev pri nekaterih panogah presenetljiva. Izstopata panoga Energetika, ki bi jo pričakovali med slabše ocenjenimi in panoga Nepremičnine, ki dosega najslabše ocene. Za ostale panoge je razvrstitev precej sorazmerna s tveganjem.

Za boljše razumevanje teh rezultatov preverim še točkovanje za vsako od mer ESG ocene posebej. Rezultati so predstavljeni v tabeli 10, kjer so panoge razvrščene glede na povprečno ESG oceno. Vsaka celica predstavlja število točk, ki jo je panoga dobila pri posamezni kategoriji. Najvišje možno število točk je bilo 100 in najnižje 1.

Slika 33: Skupna povprečna ESG ocena za vsako GICS panogo



Vir: lastno delo.

Tabela 10: Točke vsake panoge glede na mero ESG ocene

Panoga	Število ESG povedi	Odstotek ESG povedi	Povprečni sentiment	Povprečni sentiment brez nevt.	ESG sentiment	ESG sentiment brez nevt.
Energetika	58.0	82.65	59.15	54.25	53.05	55.75
Zdravstvo	68.70	60.55	49.09	50.77	55.90	58.57
Informacijska tehnologija	52.98	64.94	47.14	47.01	62.45	59.56
Nujna potrošnja	56.03	41.2	58.37	60.0	49.83	50.93
Komunikacijske storitve	50.73	64.11	52.97	53.11	40.88	39.75
Industrija	38.76	42.76	55.95	54.875	50.0	49.51
Surovine	50.83	43.46	46.46	46.37	49.06	53.17
Javne storitve	74.6	64.9	31.8	30.1	37.4	37.0
Finance	52.44	37.98	44.74	43.1	42.82	43.12
Ne nujna potrošnja	36.17	32.66	50.0	51.58	48.05	45.68
Nepremičnine	13.6	12.7	57.2	58.4	31.0	34.3

Vir: lastno delo.

Pregled zgornje tabele 10 ponudi več zanimivih ugotovitev. Panoga Energetika močno izstopa po odstotku ESG povedi v poročilu, kar nakazuje na to, da je velik del letnega poročila posvečen ESG temam, kar je skladno z velikim ESG tveganjem. Najnižji ESG odstotek je izračunan za panogo Nepremičnine, ki ima enega od nižjih ESG tveganj. Nekoliko v nasprotju s to ugotovitvijo je drugi najvišji odstotek ESG povedi za panogo IT, ki ima nizko ESG tveganje. Naslednja ugotovitev temelji na velikih razlikah (več kot 5 točk) med oceno skupnega povprečnega sentimenta in oceno ESG sentimenta. Opazimo lahko, da je pri več panogah, kjer je prisotno višje ESG tveganje, ocena skupnega sentimenta višja od ocene ESG sentimenta (primeri: Energetika, Nujna potrošnja in Industrija). Obratno, torej da je ocena ESG sentimenta občutno višja od ocene skupnega sentimenta, velja samo za dve panogi, to sta Zdravstvo in Informacijska tehnologija, ki imata najnižje ESG tveganje. Opaženo krši panoga Nepremičnin, ki ima kljub nižjemu ESG tveganju najnižjo oceno ESG sentimenta.

Skupna ESG ocena poskuša zajeti čim več mer, ki jih ponujajo metode analiza teksta uporabljene v tej raziskavi. Zanimalo me je, ali je morda še bolj skladna z ESG tveganjem ESG ocena omejena na izključno meri ESG sentimenta. V spodnji tabeli 11 so panoge razvrščene glede na sentiment v ESG povedih in sentiment v ESG povedih z izključitvijo nevtralnih povedi.

*Tabela 11: Panoge razvrščene po oceni ESG sentimenta*

<b>Panoga</b>	<b>ESG sentiment</b>	<b>ESG sentiment brez nevt.</b>
Informacijska tehnologija	62.45	59.56
Zdravstvo	55.90	58.57
Energetika	53.05	55.75
Surovine	49.06	53.17
Nujna potrošnja	49.83	50.93
Industrija	50.0	49.51
Ne nujna potrošnja	49.06	45.68

se nadaljuje

Tabela 12: Panoge razvrščene po oceni ESG sentimenta (nad.)

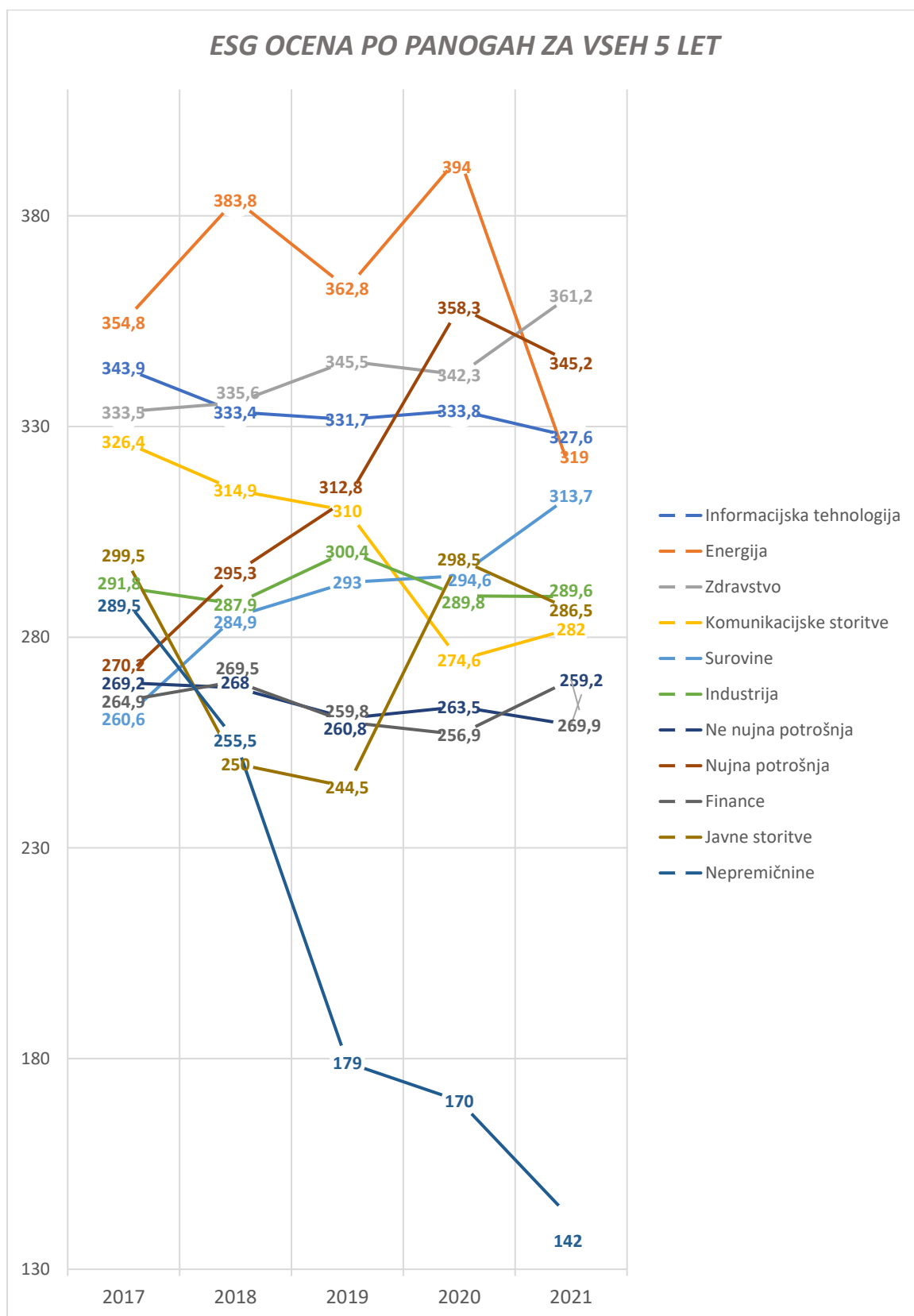
Panoga	ESG sentiment	ESG sentiment brez nevt.
Finance	42.82	43.12
Komunikacijske storitve	40.88	39.75
Javne storitve	37.4	37.0
Nepremičnine	31.0	34.3

Vir: lastno delo.

Na tej lestvici sta na prvih dveh mestih panogi Zdravstvo in IT, ki imata tudi nizko ESG tveganje, medtem ko sta na zadnjih dveh mestih panogi Nepremičnine z nizkim ESG tveganjem ter panoga Javne storitve, ki ima visoko ESG tveganje. Panoga Energetika, ki ima zelo visoko ESG tveganje, sicer izgubi dve poziciji, a je še vedno na visokem tretjem mestu.

Za še boljše razumevanje ESG ocene, so na spodnji sliki 34 ESG ocene prikazane za vsako leto posebej. To omogoča analizo trenda in variabilnosti ocene. Hitro je moč opaziti, da panoga Nepremičnin močno izstopa in ima visoko variabilnost. To lahko razložimo z dejstvom, da ima ta panoga samo dva predstavnika. Podobno velja tudi za panogo Energetika (4 predstavniki), Javne storitve (2 predstavnika) in Nujna potrošnja (6 predstavnikov). Pri ocenah panog z več predstavniki (recimo IT - 16, Zdravstvo - 15 in Industrija - 16) je variabilnost dosti nižja. Ta podatek morda razloži dejstvo, da ESG ocena panog Nepremičnin in Energetike ni skladna s pripadajočim ESG tveganjem, saj imamo opravka z majhnim številom podjetij.

Slika 34: ESG ocena za vsako panogo po analiziranih letih



Vir: lastno delo.



## 4.5 Diskusija

Izsledke empiričnega dela svojega magistrskega dela primerjam z ugotovitvami sorodne literature, ki v veliki večini zaključijo, da obstaja pozitivna povezava med CSP in CFP. Prvo postavljeno raziskovalno vprašanje se nanaša na domnevo o pozitivni povezavi med uspešnim ESG delovanjem podjetij ali uspešnim obvladovanjem ESG izzivov in finančno uspešnostjo. Uspešno ESG delovanje v magistrskem delu izrazim z ESG oceno, ki je sestavljena iz šestih različnih mer, te pa so pridobljene s pomočjo računalniške analize teksta. Za raziskovanje povezave med ESG in finančno uspešnostjo najprej podjetja razdelim v dve skupini glede na povprečno tržno kapitalizacijo. Skupino večjih podjetij po tej delitvi lahko smatramo za finančno bolj uspešno. To potrjuje tudi primerjava kazalnikov izračunanih na podlagi finančnih izkazov in tržnih podatkov. Ob primerjavi ESG uspešnosti teh dveh skupin lahko zaključim, da skupina finančno bolj uspešnih dosega povprečno višje ESG ocene. To velja za prav vse analizirane mere ESG ocene. Takšen rezultat je spodbuden, če želimo pokazati, da obstaja pozitivna povezava med CSP in CFP, vendar pa je samo na podlagi razlike v ESG ocenah nemogoče govoriti o vplivu ESG na finančno uspešnost. Večja podjetja razpolagajo z več sredstvi, kar omogoča tudi višje vložke v CSP in posledično višji pozitivni vpliv na CFP. O tem pišejo Nollet, Filis in Mitrokostas (2016), ki zaključijo, da je za pozitivno povezavo potrebno preseči določeno višino vlaganj.

V naslednjem koraku empirične analize nadaljujem z delitvijo 100 izbranih podjetij glede na povprečno ESG oceno. Pri tej delitvi me je najprej zanimalo, kako dobro s tem spremenim delitev glede na prvotni skupini. Lahko bi se namreč zgodilo, da bi skupino ESG najuspešnejših v zelo veliki večini napolnila večja podjetja po tržni kapitalizaciji. To se ne zgodi v veliki meri, saj je v skupini 50 ESG najbolj uspešnih 31 podjetij iz skupine največjih po tržni kapitalizaciji in 19 podjetij iz skupine najmanjših. To sprejemem kot dobro delitev in nadaljujem s primerjavo definiranih kazalnikov. Zanimalo me je predvsem, ali takšna delitev še vedno ohrani ali morda celo poveča razlike v primerjavi povprečnih vrednosti finančnih kazalnikov. Po primerjavi ugotovim, da skupina ESG uspešnejših dosega povprečno boljše finančne rezultate z izjemo kazalnika rast prihodkov. Ta ugotovitev je zanimiva, saj je v skupini 50 ESG najuspešnejših 19 podjetij, ki so po tržni kapitalizaciji med najmanjšimi 50 v indeksu S&P 500. Razlike so manjše kot pri prvotni delitvi podjetij, kljub temu pa lahko to ugotovitev ponovno sprejmemo kot pozitivno, če želimo pokazati pozitivno povezavo med ESG in CFP. Še posebej zanimiva je primerjava teh dveh skupin po kazalnikih izračunanih na podlagi tržnih vrednosti. Tu smo namreč lahko opazili, da ne moremo govoriti o prednosti katerekoli skupine pri povprečni letni donosnosti delnic, je pa vidna razlika pri standardnem odklonu donosnosti in varianci torej volatilnosti. Tu ima prednost skupina ESG uspešnejših. To je v skladu z ugotovitvami Ashwin Kumar in drugi (2016), ki zaključijo, da obstaja povezava med ESG uspešnostjo in nižjo volatilnostjo delnic.

Pri obeh primerjavah skupin sem bil pozoren na trend vrednosti posameznih kazalnikov skozi vseh 5 analiziranih let. Pri veliki večini primerjav sta si trenda obeh skupin zelo podobna, iz česar lahko sklepam, da so ne glede na dodeljeno skupino bodisi po tržni

kapitalizaciji bodisi po povprečni ESG oceni podjetja v enaki ali zelo podobni meri podvržena zunanjim vplivom. Tako ne morem govoriti o morebitni večji odpornosti na vplive trga katerekoli od analiziranih skupin.

Raziskovanje povezave med ESG oceno in finančno uspešnostjo nadaljujem z izračunom korelacijske matrike med ESG oceno in njenimi posameznimi merami ter kazalniki, ki so izračunani na podlagi finančnih izkazov in tržnih podatkov. Pri pregledu povprečnih korelacijskih koeficientov najprej opazim, da je višjih negativnih vrednosti zelo malo. Negativne vrednosti so višje od 0,1 le pri kazalniku ROA (povezava s številom in odstotkom ESG povedi v poročilu) ter kazalniku marža prostega denarnega toka (povezava z obema merama sentimenta celotnega poročila). Z vidika ESG in CFP povezave lahko kot negativno smatramo povprečno korelacijo med višjim odstotkom ESG povedi in višjim nihanjem donosnosti delnic, ki je izraženo z odklonom in varianco (povprečni korelacijski koeficient je 0,35). Nasprotno lahko opazimo pri obeh merah sentimenta v celotnem poročilu. Tu je namreč povprečni korelacijski koeficient z volatilnostjo negativen ( $-0,39$  in  $-0,44$ ). To nakazuje na zmerno močno korelacijo med bolj pozitivnim razpoloženjem v poročilu in nižjim nihanjem donosnosti delnic. To je skladno z ugotovitvijo iz primerjave skupin podjetij razdeljenih po ESG uspešnosti, kjer ESG bolj uspešna podjetja dosegajo nižje vrednosti standardnega odklona in variance donosnosti delnic. Pozitivnih povprečnih korelacij je v tabeli 7 možno opaziti več. Najbolj pogoste so pri kazalnikih rast prihodkov, bruto marža in neto marža. Vse omenjene vrednosti so še vedno šibke z vidika moči korelacije, vendar pa jih kljub temu lahko smatramo kot pozitivne, če želimo pokazati, da obstaja pozitivna ESG-CFP povezava.

Preverbe časovnega zamika med povezavo ESG in CSP ne izvedem, saj vsak letni zamik znotraj pridobljenih podatkov močno zmanjša velikost vzorca in s tem negativno vpliva na verodostojnost rezultatov.

Po obeh primerjavah skupin podjetij lahko tako zaključim, da podjetja z večjo povprečno tržno kapitalizacijo dosegajo višje ESG ocene. Nadalje ugotavljam, da podjetja z višjimi povprečnimi ESG ocenami dosegajo višjo finančno uspešnost izraženo z definiranimi kazalniki. Po izračunu korelacijskih koeficientov med ESG oceno in njenimi merami ter definiranimi kazalniki finančne uspešnosti lahko zaključim, da prevladujejo pozitivne vrednosti povprečnih korelacijskih koeficientov, ki nakazujejo na šibko korelacijo med višjo ESG uspešnostjo in višjo finančno uspešnostjo. Te ugotovitve so skladne z ugotovitvami sorodne literature. Na prvo definirano raziskovalno vprašanje lahko tako odgovorim pritrdilno, torej da obstaja pozitivna povezava med ESG oceno oziroma posameznimi merami ESG ocene in finančno uspešnostjo.

Tretji del empiričnega raziskovanja se osredotoča na primerjavo gospodarskih panog in preverja ugotovitve sorodne literature, da med posameznimi panogami obstajajo pomembne razlike v dojemanju dejavnikov ESG in ESG tveganjem. Raziskovalni vprašanji, ki sem jih raziskoval v tem sklopu sta, ali obstajajo pomembne razlike med ESG ocenami glede na

gospodarsko panogo in ali podjetja, ki pripadajo panogi z višjim ESG tveganjem dosegajo povprečno nižje ESG ocene. Ugotovim, da pri posameznih merah ESG obstajajo pomembne razlike med panogami. Torej da podjetja določenih panog dosegajo višje ocene pri nekaterih merah ESG ocene. Ta razlika je najbolj očitna pri povprečnem številu in odstotku ESG povedi v letnih poročilih. Prav tako lahko ugotovim, da podjetja z nižjim ESG tveganjem ocenjenim glede na podatke S&P Global, dosegajo višje ocene ESG sentimenta v primerjavi z ocenami sentimenta celotnega poročila. Zapisano drugače, deli poročila, ki se navezujejo na dejavnike ESG, so bolj pozitivno naravnana v primerjavi z razpoloženjem celotnega poročila, relativno glede na povprečne vrednosti teh mer vseh analiziranih podjetij. Ugotovitve primerjave panog so omejene z dejstvom, da raziskovane panoge niso enakovredno zastopane in so v različni meri odvisne od ocen posameznega pripadajočega podjetja. Tudi pri tem delu lahko zaključim, da se ugotovitve magistrskega dela skladajo z ugotovitvami sorodne literature.

#### **4.6 Omejitve, izboljšave in priporočila za nadaljnje raziskave**

Glavne omejitve raziskave magistrskega dela izhajajo iz natančnosti računalniških metod, ki niso stoodstotne. To pomeni, da pri vsakem koraku obdelave izgubim nekaj informacij ali jih ne klasificiram ustrezno.

Prvi korak obdelave podatkov je ekstrakcija povedi iz letnih poročil podjetij. Kljub temu, da vseh 500 analiziranih letnih poročil sledi formatu 10K, so končne oblike besedil z vidika računalniške obdelave nekoliko različne. To pomeni, da je bilo potrebno algoritem za prepoznavanje povedi v manjši meri prilagoditi glede na določeno vhodno besedilo. Prav tako je zelo težko izmeriti natančnost tega algoritma oziroma izračunati kolikšen odstotek dejansko prisotnih povedi je algoritem uspešno prepoznal in shranil za obdelavo. Ta postopek bi bil lahko izboljššan z analizo večjega števila posameznih poročil in ugotavljanjem, do kakšnih napak prihaja pri ekstrakciji povedi. Takšen postopek bi bil zelo zamuden in je za veliko količino podatkov, kar 500 letnih poročil zagotovo je, skorajda neuporaben.

Drugi del pridobivanja podatkov temelji na prepoznavanju povedi, ki se navezujejo na dejavnike ESG. Na manjši podmnožici izmerjena natančnost ESG modela, ki napove, ali se poved nanaša na ESG teme ali ne, je okoli 70%. V tem postopku tako lahko pride do napake, da je določena poved napačno klasificirana kot ESG poved (napaka tipa I), ali da je poved napak klasificirana kot NE-ESG poved (napaka tipa II). Do podobnih napak kot pri prepoznavanju ESG povedi lahko pride pri izhodnih vrednostih modela za napovedovanje sentimenta ali razpoloženja. Povedi je lahko tako dodeljeno napačno razpoloženje. Oba modela sta odvisna od uporabljenih algoritmov strojnega učenja in metod analize teksta. Hiter razvoj na tem področju omogoča vedno bolj natančno napovedovanje in klasifikacijo. Uporabljen model BERT trenutno predstavlja najsodobnejšo rešitev (angl. state of the art) za analizo sentimenta, a lahko zagotovo v bližnji prihodnosti pričakujemo še natančnejše

metode. Vse bolj napredne računalniške rešitve pa običajno zahtevajo tudi vse višjo računsko zmogljivost strojne opreme. Vse računalniške rešitve uporabljene tekom izdelave magistrskega dela so bile izvedene na osebnem računalniku, tako da je bila izvedba nekaterih algoritmov precej dolgotrajna.

Naslednja omejitev raziskave izhaja iz izbora analiziranih podjetij. Na ta seznam sem vključil 100 ameriških podjetij iz skupine S&P 500. Povečanje tega seznama bi lahko pomenilo zanesljivejše rezultate. Ta omejitev je še posebej očitna pri analizi podjetij po gospodarskih panogah, kjer so bile panoge neenakovredno zastopane. Zelo podobno bi pozitivno na rezultate lahko vplival povečan obseg analiziranih finančnih let. V magistrskemu delu se omejim na zadnjih 5 let. Večje število podjetij ali večji obseg let bi pomenil povečano število analiziranih letnih poročil. Prav tako bi daljši časovni razpon podatkov omogočal analiziranje ESG-CFP povezave z več letnim zamikom.

Ugotovitve magistrskega dela usmerjajo nadaljnje raziskovanje predvsem v smeri analize povezave med določenimi vidiki ESG in posameznimi vidiki finančne uspešnosti. Glede na ugotovitve sorodne literature bi bilo smiselno to povezavo raziskovati skozi daljše obdobje. V svoji raziskavi se osredotočim na ameriška podjetja. Raziskavo bi v podobnem okviru bilo mogoče izvesti za podjetja celega sveta, kar bi omogočalo medsebojno primerjavo držav, kontinentov in politično-ekonomskih zvez. Na ta način bi lahko bolje raziskali ugotovitev nekaterih študij, da obstajajo pomembne razlike pri povezavi med CSP in CFP v podjetjih držav v razvoju.

## **SKLEP**

Magistrsko delo je nastalo z namenom doprinosa k boljšemu razumevanju povezave med finančno uspešnostjo poslovanja podjetij in njihovo usmerjenostjo v trajnostno poslovanje. V začetku postavim 8 ciljev, ki jih želim z magistrskim delom doseči.

V prvem poglavju opišem dejavnike ESG in celostno predstavim ključne ugotovitve sorodne literature. Ugotovitve najprej uporabim za definicijo treh raziskovalnih vprašanj in nato za primerjavo z izsledki mojega raziskovalnega dela. Drugo poglavje se osredotoča na predstavitev računalniških tehnologij za pridobivanje informacij in analizo teksta. V poglavju podrobno opišem metode in tehnologije, ki jih običajno uporabimo pri projektih rudarjenja besedil. Predstavitev metod si sledi v vrstnem redu, kot sem ga uporabil tudi sam tekom izdelave magistrskega dela. To bralcu omogoča strukturiran pregled tematike kot tudi uporabo metod za namene svojega dela.

Med razvojnim delom raziskave najprej razvijem algoritem za pridobivanje podatkov iz spleta. V drugem delu razvijanja napišem algoritem za pridobivanje informacij iz delno strukturiranih besedil oziroma letnih poročil podjetij. V naslednjem delu razvijem model za prepoznavanje povedi, ki se nanašajo na dejavnike ESG in model za oceno prevladujočega razpoloženja v povedih. Delovanje obeh modelov evalviram na podmnožici podatkov. Z

uporabo razvitih algoritmov za vsako podjetje pridobim ESG oceno in pripadajočih 6 mer ESG uspešnosti.

V empiričnem delu raziskave z uporabo statističnih metod raziščem povezavo med pridobljenimi merami ESG in posameznimi vidiki finančne uspešnosti, ki so izraženi s kazalniki, ki so izračunani na podlagi finančnih izkazov in tržnih podatkov. V zaključnem delu rezultatov raziščem še povezavo med ESG oceno in ESG tveganjem, ki so mu podvržena podjetja. Zadnji del četrtega poglavja je posvečen povzetku ključnih ugotovitev in primerjavo z ugotovitvami sorodnih študij. Prav tako predstavim omejitve magistrskega dela in predlagam nadaljnje usmeritve raziskovanja povezave med CSP in CFP.

Verjamem, da magistrsko delo dosega vseh 8 zastavljenih ciljev in v celoti izpolnjuje svoj namen. Predstavlja še en način uporabe vedno bolj zmogljivih računalniških tehnologij za raziskovanje družbenega vprašanja, ki je odprto že več desetletij. Delo uspešno združuje področja računalništva, ekonomije in lingvistike ter tako sledi smernicam interdisciplinarnega raziskovanja, ki sintetizira ideje različnih disciplin in povezuje različne karakteristike področij.

Področje obdelave naravnega jezika vedno bolj pridobiva zanimanje izven računalništva in informatike. Vedno močnejše je zavedanje, da lahko s premišljeno in inovativno uporabo te tehnologije izboljšujemo obstoječe procese in s tem pozitivno vplivamo na razvoj družbe. Analiza naravnega jezika že v osnovi združuje oziroma zahteva znanja različnih področij. Vse večja potreba po naprednih NLP rešitvah predstavlja nov izziv, kako uspešno združiti in uporabiti najnovejša dognanja vseh vključenih disciplin in ved. Verjamem, da magistrsko delo predstavlja dober primer takšnega združevanja.

S področja računalništva, informatike in lingvistike v magistrskem delu opišem in uporabim različne računalniške tehnologije, ki nam omogočajo, da iz količine podatkov, ki je sicer zaradi svoje velikosti ne bi mogli preučiti, pridobimo podatke, ki povečajo naše znanje in izboljšajo zmožnost odločanja. Na ta način pridobljene podatke nato z uporabo statističnih metod pretvorim v znanje na področju ekonomije.

Zahtevnost uporabe naprednih računalniških metod lahko predstavlja oviro pri reševanju NLP problemov. Način obdelave podatkov, ki ga predstavim v tem delu, je moč prevesti na drugačno vprašanje. Natančen opis vseh postopkov in uporabljenih rešitev je tako koristen tudi izven problematike raziskovane v magistrskem delu.

Ugotovitve magistrskega dela predstavljajo še en vidik povezave med CSP in CFP. Glavni doprinos dela vidim v inovativnem načinu izračuna oziroma ocene ESG delovanja posameznega podjetja. Po pregledu literature ugotovim, da vprašanje, kako izmeriti ESG uspešnost nima enega samega pravilnega odgovora. Tako lahko sklepam, da raziskovanje novih načinov, kako odgovoriti na to vprašanje, lahko v veliki meri koristi boljšemu razumevanju povezave med trajnostno zavezanostjo in družbeno odgovornostjo ter finančno uspešnostjo.

## LITERATURA IN VIRI

1. Admirals. (2022). *Gospodarski sektorji – 11 delniških sektorjev za investiranje*. Pridobljeno 25. junija 2022 iz <https://admiralmarkets.com/si/education/articles/shares/gospodarski-sektorji>
2. Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing & Management*, 39(1), 45-65.
3. Alammari, J. (2021). *The Illustrated BERT, ELMO, and co. (How NLP Cracked Transfer Learning)*. Pridobljeno 1. junija 2022 iz <http://jalammari.github.io/illustrated-bert/>
4. Alberti, C., Lee, K. & Collins, M. (2019). A bert baseline for natural questions. *arXiv preprint arXiv:1901.08634*.
5. Ashwin Kumar, N. C., Smith, C., Badis, L., Wang, N., Ambrosy, P. & Tavares, R. (2016). ESG factors and risk-adjusted performance: A new quantitative model. *Journal of Sustainable Finance & Investment*, 6(4), 292-300.
6. Balakrishnan, V. & Lloyd-Yemoh, E. (2014). Stemming and lemmatization: a comparison of retrieval performances. *Lecture Notes on Software Engineering*. 2. 262-267. 10.7763/LNSE.2014.V2.134.
7. Barnett, M. L. & Salomon, R. M. (2012). Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic management journal*, 33(11), 1304-1320.
8. Boaventura, J. M. G., Silva, R. S. D. & Bandeira-de-Mello, R. (2012). Performance Financeira Corporativa e Performance Social Corporativa: desenvolvimento metodológico e contribuição teórica dos estudos empíricos. *Revista Contabilidade & Finanças*, 23, 232-245.
9. Brundtland, G. H. (1987). Our common future—Call for action. *Environmental Conservation*, 14(4), 291-294.
10. Busch, T. & Friede, G. (2018). The robustness of the corporate social and financial performance relation: A second-order meta-analysis. *Corporate Social Responsibility and Environmental Management*, 25(4), 583-608.
11. Cambria, E. & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine*, 9(2), 48-57.
12. Corey, D. M., Dunlap, W. P. & Burke, M. J. (1998). Averaging correlations: Expected values and bias in combined Pearson rs and Fisher's z transformations. *The Journal of general psychology*, 125(3), 245-261.
13. Crilly, D., Zollo, M. & Hansen, M. T. (2012). Faking it or muddling through? Understanding decoupling in response to stakeholder pressures. *Academy of Management Journal*, 55(6), 1429-1448.
14. Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

15. Devlin, J. & Chang, M. W. (2018). Open sourcing BERT: State-of-the-art pre-training for natural language processing. *Google AI Blog*, 2.
16. Dodds, L. (2020). *Do Data Scientists Spend 80 Of Their Time Cleaning Data Turns Out No* [objava na blogu] . Pridobljeno 21. maja 2022 iz <https://blog.ldodds.com/2020/01/31/do-data-scientists-spend-80-of-their-time-cleaning-data-turns-out-no/>
17. Dofleitner, G., Halbritter, G. & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility—An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, 16(7), 450-466.
18. Dossi, A. & Patelli, L. (2010). You learn from what you measure: financial and non-financial performance measures in multinational companies. *Long Range Planning*, 43(4), 498-526.
19. Eccles, R. G., Krzus, M. P., Rogers, J. & Serafeim, G. (2012). The need for sector-specific materiality and sustainability reporting standards. *Journal of applied corporate finance*, 24(2), 65-71.
20. Egorova, A. A., Grishunin, S. V. & Karminsky, A. M. (2022). The Impact of ESG factors on the performance of Information Technology Companies. *Procedia Computer Science*, 199, 339-345.
21. European Commission. (2001). Promoting a European Framework for Corporate Social Responsibility. *COM(2001) 366*.
22. Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4), 507-521.
23. Fitzpatrick, G., Neilan, J. & Reilly, P. (2020). *Time to rethink the S in ESG*. Pridobljeno 25. maja 2022 iz <https://corpgov.law.harvard.edu/2020/06/28/time-to-rethink-the-s-in-esg/>
24. Fornasari, F. (2020). Knowledge and power in measuring the sustainable corporation: stock exchanges as regulators of ESG factors disclosure. *Washington University Global Studies Law Review*, 19, 167.
25. Friede, G., Busch, T. & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233.
26. Fumo, D. (2017). *Types of Machine Learning Algorithms You Should Know*. Pridobljeno 28. maja 2022 iz <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861>
27. Garcia, A. S. & Orsato, R. J. (2020). Testing the institutional difference hypothesis: A study about environmental, social, governance, and financial performance. *Business Strategy and the Environment*, 29(8), 3261-3272.
28. Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc.

29. Hübel, B. & Scholz, H. (2020). Integrating sustainability risks in asset management: The role of ESG exposures and ESG ratings. *Journal of Asset Management*, 21(1), 52-69.
30. Investopedia. (2022). *Environmental, Social and Governance (ESG) Principles and Criteria*. Pridobljeno 6. junija 2022 iz <https://www.investopedia.com/terms/e/environmental-social-and-governance-esg-criteria.asp>
31. Jordan, M. I. & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
32. Kannan, S., Gurusamy, V., Vijayarani, S., Ilamathi, J., Nithya, M., Kannan, S. & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.
33. Karimi, A., Rossi, L. & Prati, A. (2020). Improving bert performance for aspect-based sentiment analysis. *arXiv preprint arXiv:2010.11731*.
34. Kiss, T. & Strunk, J. (2006). Unsupervised multilingual sentence boundary detection. *Computational linguistics*, 32(4), 485-525.
35. La Torre, M., Mango, F., Cafaro, A. & Leo, S. (2020). Does the esg index affect stock return? evidence from the eurostoxx50. *Sustainability*, 12(16), 6387.
36. Leemeta. (2021). *Kakšna je razlika med stavkom in povedjo?* Pridobljeno 28. maja 2022 iz <https://www.leemeta.si/blog/slovnica/kaksna-je-razlika-med-stavkom-povedjo>
37. Loper, E. & Bird, S. (2002). Nltk: The natural language toolkit. *arXiv preprint cs/0205028*.
38. Malo, P., Sinha, A., Korhonen, P., Wallenius, J. & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology*, 65(4), 782-796.
39. Margolis, J. D., Elfenbein, H. A. & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance. *SSRN Electronic Journal*. 10.2139/ssrn.1866371.
40. Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
41. Mikolov, T., Le, Q. V. & Sutskever, I. (2013). Exploiting similarities among languages for machine translation. *arXiv preprint arXiv:1309.4168*.
42. Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.
43. Nollet, J., Filis, G. & Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400-407.
44. Orlitzky, M., Schmidt, F. L. & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization studies*, 24(3), 403-441.
45. Pástor, L., Stambaugh, R. F. & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2), 550-571.
46. Patel, K. & Bhattacharyya, P. (2017). Towards lower bounds on number of dimensions for word embeddings. *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 31-36).



47. Porter, M. E. & Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives*, 9(4), 97-118.
48. Ratner, B. (2009). The correlation coefficient: Its values range between+ 1/- 1, or do they?. *Journal of targeting, measurement and analysis for marketing*, 17(2), 139-142.
49. Samuel, A. L. (1959). Machine learning. *The Technology Review*, 62(1), 42-45.
50. Sassen, R., Hinze, A. K. & Hardeck, I. (2016). Impact of ESG factors on firm risk in Europe. *Journal of business economics*, 86(8), 867-904.
51. Sciarelli, M., Cosimato, S., Landi, G. & Iandolo, F. (2021). Socially responsible investment strategies for the transition towards sustainable development: The importance of integrating and communicating ESG. *The TQM Journal*.
52. Shinohara, Y. (2014). A study on CSP and CFP relationships: When does CSP effect on CFP. And Which Stakeholder Is the Most Influential on CFP. *Mita Business Review*, 57, 21-45.
53. Silver, N. C. & Dunlap, W. P. (1987). Averaging correlation coefficients: should Fisher's z transformation be used? *Journal of applied psychology*, 72(1), 146.
54. Singh, A. K. & Shashi, M. (2019). Vectorization of text documents for identifying unifiable news articles. *Int. J. Adv. Comput. Sci. Appl*, 10(7).
55. S&P Global. (2019). *ESG Industry Report Card: Regulated Utilities Networks*. Pridobljeno 12. junija 2022 iz [https://www.spglobal.com/\\_media/documents/ratingsdirect\\_esgindustryreportcardregulatedutilitiesnetworks\\_41535202\\_may-17-2019.pdf](https://www.spglobal.com/_media/documents/ratingsdirect_esgindustryreportcardregulatedutilitiesnetworks_41535202_may-17-2019.pdf)
56. S&P Global. (2020). *What is the "G" in ESG?* Pridobljeno 15. junija 2022 iz <https://www.spglobal.com/en/research-insights/articles/what-is-the-g-in-esg>
57. Študenti Ekonomske fakultete. (2021). *Zaznavanje ESG vsebin v povedih in ustrezna klasifikacija* [podatkovna datoteka]. Ljubljana: Ekonomska fakulteta.
58. The Global Compact. (2004). Who cares wins: Connecting financial markets to a changing world. *New York*.
59. Trillium Asset Management. (2021). *ESG Integration & Criteria*. Pridobljeno 25. julija 2022 iz <https://www.trilliuminvest.com/esg/esg-integration-criteria>
60. Van Beurden, P. & Gössling, T. (2008). The worth of values—a literature review on the relation between corporate social and financial performance. *Journal of business ethics*, 82(2), 407-424.
61. Wilbur, W. J. & Sirotkin, K. (1992). The automatic identification of stop words. *Journal of information science*, 18(1), 45-55.
62. Xiong, J. X. (2021). The Impact of ESG Risk on Stocks. *The Journal of Impact and ESG Investing*, 2(1), 7-18.