



Stockholms
universitet



HÖGSKOLAN
I GÄVLE



Mittuniversitetet
MID SWEDEN UNIVERSITY

Hållbarhetsprediktion med Maskininlärning

eller utveckling av en

Hållbarhetsrobot

Slutrapport för projekt med referensnummer P8/18

Dr. Jan Svanberg

Dr. Tohid Ardeshiri

Dr. Isak Samsten

Professor Peter Öhman

Professor Presha E. Neidermeyer

Professor Mats Danielson

Dr. Tarek Rana

(Från högskolor och universitet huvudsakligen bidragande till projektet)

Förord

I det följande redovisas projektet *Hållbarhetsprediktion med maskininlärning*, P08/18 (tidigare under annat namn).

Tack till Länsförsäkringsbolagens Forskningsfond som finansierat projektet och till flera på Länsförsäkringar som med kunskaper sett till att projektet blivit av: Lars Höglund, Kristofer Dreiman, Alexander Elving, Mari Sparr, och sist men inte minst Peter Griepenkerl-Lööf som handgripligen alla varit inblandade och bidragit med lösningar vid en rad möten.

Melbourne den 24 maj 2021

Jan Svanberg

Innehållsförteckning

Sammanfattning	s. 3
Forskningsleverans	s. 4
Sammanfattning av studierna om ESG-betyg	s. 4
Tidskriftsartiklar i sammanfattning	s. 8
Böcker och bokkapitel	s. 10
Praktisk Leverans	s. 10
HBR med Prediktiv Modelling	s. 10
Precision och prediktiv validitet	s. 12
Betygsexempel och beskrivning av funktion	s. 13
Tolkning av betyg	s. 17
Avslutande kommentar	s. 22

Sammanfattning

Syftet med vårt projekt är att utveckla en metod att bedöma hur 'hållbara', eller, med ett begrepp som vi föredrar, 'socialt ansvarstagande' företag är. Metoden är tänkt att kunna användas av Länsförsäkringar Kapitalförvaltning (LF Kapital) som har behov av att kunna bedöma hållbarheten hos åtminstone 1500 – 3000 företag i vilka förvaltningen har investerat och många fler företag i vilka LF Kapital skulle kunna vara intresserat att investera. I hållbarhetsbedömning ingår vanligen tre områden som på engelska förkortas ESG, där E står för miljö, S för social, och G för bolagsstyrning. Kommersiella företag eller institut tillhandahåller beräknade ESG-betyg (tex Sustainalytics, MSCI, Refinitiv, Bloomberg), men forskare och institutionella kapitalförvaltare utvecklar också egna ESG-betyg eller liknande nyckeltal.

ESG-betyg är sammanfattande bedömningar av mycket omfattande begrepp. Enbart för miljökomponenten rör det sig om många viktiga typer av effekter på miljön, tex utsläpp i luft, mark och vatten, avfall, energi- och vatten- och annan resursanvändning, samt policys och liknande ansträngningar att minska hållbarhetsrisker.

Det finns omfattande hållbarhetsdata härrörande huvudsakligen från företags årsredovisningar, men också från andra källor. Mängden uppgifter skiljer sig åt mellan databaser men vi har haft tillgång till ca 100 indikatorer för var och en av komponenterna E, S och G från en stor databas. Problemet med mätning av hållbarhet är därför inte datatillgång utan kvaliteten hos data och konstruktionen av metoderna med vilka data analyseras. De ESG-betyg som framställs av betygsindustrin och som används i institutionell kapitalförvaltning har konstaterats sakna validitet (Berg et al., 2019; Chatterji et al., 2016; Chen and Delmas, 2011; Delmas et al., 2013; Semenova and Hassel, 2015; Trumpp et al., 2013). Vårt projekts syfte är att lösa det problemet.

Vi har utvecklat en ny typ av ESG-betyg med maskininlärning och med ett revolutionerande sätt att använda hållbarhetsdata. Lösningen på validitetsproblemet med ESG-betyg är utvecklad dels i fyra *empiriska studier*, dels som *en produkt bestående av programvaror som genererar valida ESG-betyg*.

Vår *forskningsleverans* är (1) *sju tidskriftsartiklar* skickade till vetenskapliga tidskrifter enligt förteckning, samt (2) *två bokkapitel* i två separata böcker utgivna på Springer förlag och (3) vårt eget *redaktörskap för Springer 'Handbook of Big Data and Analytics in Accounting and Auditing'*, omfattande 60 kapitel, tillsammans med två kollegor på RMIT University professor Alan Lowe och Dr. Tarek Rana. Jan har också varit Research Fellow vid RMIT University (The Royal Melbourne Institute of Technology) mellan 1 september 2019 och 30 april 2021. Universitetet har internationellt anseende som ett av de 20 bästa universiteten i världen för universitet yngre än 50 år enligt QS university ranking. RMIT University är nyligen ansett som nummer tre i världen hos The Times Higher Education Impact Rankings som är en bedömning av universitetens prestationer mot FN:s 17 hållbarhetsmål. P8/18 presenterades vid det stora forskningsseminariet vid RMIT University, School of Accounting i februari 2020.

Vår *praktiska leverans* definieras av det som i den ursprungliga ansökan beskrivs med uttrycket en undersökning av *"förutsättningarna för ett system för kontinuerlig internrevision"* vilket, med den riktningförändring som gjorts ämnesmässigt sedan Länsförsäkringar inte kunde finna tillgång till data för

det ursprungliga projektet, blir en undersökning av ”förutsättningarna för ett system för analys av hållbara investeringar”, som Peter Griepkerl-Lööf döpt till ’Hållbarhetsrobot’.

Med beaktande av att syftet med P8/18 är att utvinna kunskap har vi parallellt med forskningsprojektet utvecklat en produkt som vi fortsättningsvis kallar Hållbarhetsroboten (HBR). Hållbarhetsroboten har utvecklats helt och hållet med data från databaser som Högskolan i Gävle betalar för och tillhandahåller via högskolebiblioteket. Ingen data av karaktären affärshemligheter från LF Kapital har använts i utvecklingen.

HBR bedömer hållbarhet i dimensionerna E, S och G för 2517 företag, och med kommande arbete under hösten 2021 då ”förutsättningarna för ett system för analys av hållbara investeringar” kan omsättas i praktisk produktutveckling, och med insatser av professionella programmerare, bedömer hållbarhetsanalyssystemet enligt vår prognos hållbarhet för ca 10 000 stora företag motsvarande över 80 procent av världens aktiekapital och får, förutom ett revolutionerande användargränssnitt, en lång rad användbara funktioner. HBR är den enda nu existerande metoden att bedöma hållbarhet som har (prediktiv) validitet. Validitet är ett viktigt framsteg inte bara forskningsmässigt utan också praktiskt. Vi finner därför att ”förutsättningarna för ett system för analys av hållbara investeringar” föreligger i allra högsta grad och att den metod som HBR förkroppsligar är en plattform för fortsatt produktutveckling.

Forskningsleverans

Sammanfattning av studierna om ESG-betyg

Bakgrund: ESG-betyg ges ofta på tre områden separat, E, S och G. ESG-betyg kan också beräknas som ett holistiskt omdöme som sammanfattar de tre prestationsområdena. När miljöbetyg beräknas separat kallar vi dem miljöbetyg eller E-betyg, och motsvarande gäller för S, S-betyg eller sociala betyg och G, G-betyg eller bolagsstyrningsbetyg. Dessa betyg är sammanfattande bedömningar av i vilken mån ett företag tar sitt hållbarhetsansvar i absoluta termer eller i relation till andra företag. Vi fokuserar den fortsatta beskrivningen av forskningsleverans på miljöbetyg och hela ESG-betyg men det vi skriver om miljöbetyg fungerar på ett liknande sätt för de två andra betygen.

Med tanke på att ESG-betyg används i stor utsträckning av institutionella förvaltare är det förvånande att ingen undersökning har lyckats visa att något enda ESG-betyg har validitet som mått på hållbarhet. Det är också oklart vad enskilda leverantörer av ESG-betyg menar med hållbarhet även om de flesta leverantörer ringar in ungefär samma områden av företagsegenskaper som relevanta (Berg et al., 2019). Brister och oklarheter om vad betygen egentligen mäter och hur tillförlitliga de är har lett till en växande misstro bland institutionella förvaltare. I förhållande till ESG-betyg genomgår institutionella förvaltare enligt den årliga undersökningen *Rate the raters "the great disillusionment – and getting more dissatisfied the more they look under the hood"* (Wong and Petroy, 2020, p. 30). Brister hos ESG-betyg som beslutsunderlag vid hållbara investeringar har identifierats som en utmaning för portföljsammansättning (Ballesterio et al., 2012; Oikonomou et al., 2018; Utz, 2019) och studier som

undersökt validiteten hos ESG-betyg har rapporterat övertygande bevis för att validitet saknas (cf. Berg et al., 2019; Chatterji et al., 2016; Semenova and Hassel, 2015) varför det kan ifrågasättas vad ESG-betyg egentligen mäter (cf. Chatterji and Levine, 2006; Chen and Delmas, 2011; Delmas and Blass, 2010; Delmas et al., 2013; Trumpp et al., 2013). Även om engagemanget för hållbara investeringar alltså är högt bland institutionella förvaltare kan metoderna de använder att identifiera hållbara företag i värsta fall vara så oprecisa att resultatet av ansträngningarna är ineffektiva (Hartzmark och Sussman, 2019). Problemet har kritiserats skarpast av Chatterji¹ et al. (2016) som i en stor studie publicerad i världens högst ansedda tidskrift i operationsanalys avslutar sin undersökning av validiteten hos kommersiellt ledande leverantörers ESG-betyg med: “[The] low convergent validity we report implies that the results of prior academic studies using these metrics should be reassessed. Thus, we urge users to provide evidence that the ratings are sufficiently valid for their intended purposes” (p. 1608). En annan aspekt av hur ESG-betygens brister kan ha skadliga effekter på affärspraxis är att företagsledare i sin strävan efter bättre hållbarhet och anseende kan söka efterleva ESG-betyg vilka till följd av sina brister riskerar leda företagets hållbarhetsansträngningar i fel riktning. Vi ser tre konstruktionsfel med traditionella ESG-betyg och diskuterar dessa i korthet för miljöbetyg:

Det första problemet med konventionella ESG-betyg är att de är beräknade med linjära modeller. Berg, Köbel & Rigobon (2019) fann att sex ledande leverantörers ESG-betyg kan reproduceras med linjära modeller. Chen & Delmas (2011) fann att ESG-betyg huvudsakligen är aritmetiska, ad hoc- eller likaviktade medelvärden av hållbarhetsindikatorer. Miljöprestation eller miljöansvarstagande är dock ett heterogent, flerdimensionellt och icke-linjärt begrepp (Oikonomou, Platanakis, & Sutcliffe, 2018). Enskilda aspekter av miljöprestation interagerar sannolikt med varandra (effekten av en indikator är beroende av om en annan indikator har högt eller lågt värde). En policy för avfallsminskning interagerar med mängden avfall: Effekten av policyn att minska på avfall bör vara obefintlig om företaget inte har något avfall men kan vara stor om företaget har mycket avfall. Linjära modeller som ignorerar de många potentiella interaktionerna saknar kapacitet att representera denna komplexitet. De bedömer alltså policyn för avfallsminskning som lika viktig för miljöbetyget oavsett om företaget har mycket avfall eller (nästan) inget avfall alls. Särskilt tydligt blir det att linjära modeller är olämpliga för miljöbetygsberäkning om vi definierar miljöprestation som företags efterlevnad av miljökrav. Utsläpp upp till den nivå som miljökraven anger är inte alls klandervärda medan utsläpp över denna nivå är olagliga eller omoraliska och därför klandervärda. Ett miljöbetyg som tar sikte på efterlevnad av miljökrav bör reagera med en sänkning av miljöbetyget när ett företags utsläpp överskrider ett miljökrav. Betraktat som efterlevnad av miljökrav har miljöprestation, och miljöbetyget, formen av ett trappsteg. Linjära betygsmodeller bör därför vara olämpliga eftersom sådana, så kallade *high-bias models* (Lundberg et al., 2019), är behäftade med systematiska fel även vid moderat icke-linjär data.

Det andra problemet med traditionella ESG-betyg är att godtyckliga indikatorvikter ger godtyckliga ESG-betyg (Callan och Thomas, 2009). Om vi gör det orealistiska antagandet att vi istället för ca 100 indikatorer på miljöprestation skulle ha bara tre: Koldioxid, avfall, och vattenanvändning.

¹ Aaron Chatterji, Ph.D. är docent vid University of California, Berkeley. Han har tidigare varit seniorekonom vid Vita husets råd för ekonomiska rådgivare (CEA).

Beroende på hur stor vikt vår betygsmodell fäster vid var och en av dessa får vi olika miljöbetyg för ett och samma företag. Vi kommer till olika slutsatser om vilka företag som är minst miljöskadliga beroende på vilken vikt vi höjer och vilken vi sänker. Vi kan bedöma att koldioxid eller växthusgaser är mycket viktiga i energisektorn och därför bör ha störst vikt där men exakt hur mycket större än vikterna för avfall och vattenanvändning? Där har varken forskning eller finansindustrin något exakt svar trots en långvarig debatt (cf. Hillman and Keim, 2001). Vissa producenter av ESG-betyg hänvisar till att de sätter vikter på indikatorer efter hur finansiellt materiella (relevanta) de är för företagen (eller aktiernas prisutveckling) i förhållande till hur finansiellt materiella andra indikatorer är för samma företag, men forskningen om finansiella effekter av enskilda aspekter av ESG är inte tillräcklig för att fungera som underlag för bestämning av vilka vikter olika företagsbeteende bör ha vid beräkning av ESG-betyg. Leverantörer av denna typ av ESG-betyg redovisar inte vilka studier de väljer sina indikatorvikter med stöd av. Det är, enligt vår uppfattning, omöjligt att med precision skatta den relativa finansiella (framtida) effekten av en ESG-indikator i förhållande till en annan. Finansiell materialitet är dock en aspekt av hållbarhet som anses mycket relevant av finansindustrin och en hel del forskning har funnit att hållbara företag kan vara mer lönsamma än ohållbara företag och att investeringar i hållbara företag kan vara mer lönsamt för institutionella förvaltare (cf. Nofsinger et al., 2019). Beaktande av finansiell materialitet är därför viktigt och av stort intresse för vårt projekt men troligen en oframkomlig väg att utveckla en hållbarhetsmodell som med hög precision kan identifiera hållbara företag.

Det tredje problemet är att traditionella ESG-betyg är svåra eller i värsta fall omöjliga att tolka. ESG-betygsmodeller som åberopar finansiell materialitet är tyngda av svårigheten att visa att modellvikterna representerar ett objektiva mått på miljöansvar. I brist på bevis om vad modellens indikatorvikter betyder blir tolkningen en spekulering om relativ finansiell materialitet. Ett annat tolkningsproblem beror på användningen av linjära betygsmodeller: Lundberg et al. (Lundberg et al., 2019) visar i en studie av linjära modellers tolkningsbarhet som funktion av icke-linjäritet hos data att även moderat icke-linjäritet leder till att linjära modeller utpekar irrelevanta indikatorer som betydelsefulla för modellens skattning. Användning av linjära betygsmodeller leder därför till att irrelevanta hållbarhetsindikatorer felaktigt identifieras som betydelsefulla för företags ESG-betyg.

Traditionella ESG-betyg är inte bara missvisande till följd av beräkningsmodellernas konstruktion utan är också oförenliga med institutionella investerares informationsbehov. Ny forskning i den världsledande tidskriften *Journal of Corporate Finance* har visat att institutionella investerare har selektiva preferenser för ESG. Institutionella investerare undviker att investera i företag som tenderar att vara inblandade i hållbarhetskontroverser (skandaler) och är därför huvudsakligen intresserade av hållbarhetsinformation som avslöjar om företag uppfyller hållbarhetsförpliktelser som är obligatoriska, medan frivilliga åtaganden är av mindre intresse från investerarnas perspektiv (Nofsinger et al., 2019). Anledningen till investerarnas ensidiga intresse för hållbarhetsinformation som uttrycker företagets normefterlevnad är att bristande efterlevnad av juridiska eller moraliska hållbarhetskrav leder till stora *negativa effekter* som rättstvister, straff och böter, kundbojkotter och produktionsstörningar (Benabou och Tirole, 2010) medan företags frivilliga hållbarhetsarbete kostar ungefär lika mycket för företagen som de belönas av sin omgivning för att utföra det. Att göra gott i frivilliga frågor kan ha fördelar för företaget,

t.ex. lägre kapitalkostnad och lägre risk (Dyck et al., 2019) men fördelarna kompenseras helt av högre kostnader för att fullgöra handlingarna under de flesta förhållanden. Nettoeffekten är finansiellt irrelevant för institutionella investerare. Marknader reagerar starkt på dåliga ESG-nyheter men svagt på goda ESG-nyheter (Capelle-Blancard och Petit, 2019; Krüger, 2015). Krueger et al. (2020) fann att institutionella investerares tyngst vägande skäl att beakta ESG vid investeringar är att skydda det egna ryktet som ansvarsfull investerare, vilket görs genom att undvika att investera i företag som brister i socialt ansvarstagande, dvs är benägna att bli föremål för hållbarhetskontroverser. Institutionella investerare har också etiska motiv (Amir och Serafeim, 2018) men drivs framför allt av ett instrumentellt intresse av att undvika att ha kontroversbenägna företag i sina (hållbara) portföljer.

Som ett svar på dessa informationsbehov hos institutionella investerare föreslår vi en normefterlevnads-baserad ESG-metod med begreppsliga rötter i Woods (2010, 1991) ledande konceptualisering av hållbarhetsprestation (ofta kallas prestationen CSP (corporate social performance) och hållbarhetsansvaret CSR (corporate social responsibility)), och som är nära relaterad till Carrolls (1979) hållbarhetsklassiker. Woods modell kallas "*one of the most influential, helpful, parsimonious and yet comprehensive conceptualizations of CSP*" (Orlitzky et al., 2003, p. 411). Enligt Wood är definitionen av hållbarhetsprestation relaterad till legitimitet och understryker vikten av juridiskt och moraliskt ansvar som den begreppsliga grunden för ESG. Vi begränsar därför vår definition av ESG till den obligatoriska aspekten av hållbarhetsprestation där företagens hållbarhetsbeteende bedöms mot normer för miljö- och socialt ansvar samt normer för bolagsstyrning.

Vi ser hållbarhetskontroverser som uttryck för företags bristande efterlevnad av hållbarhetsförpliktelser. Kontroverser är tecken på systematiska svagheter i företag (Bernile och Jarrell, 2009; Jory et al., 2015; Kochan, 2002; Nieri och Giuliani, 2018) tex bristande efterlevnad av styrelseansvar, ingrepp i rätten för arbetstagare att organisera sig fackligt eller brott mot utsläppsregler. Vi använder databaser med fler än 10 000 kontroverser för att skilja mellan företag som har kontroverser eller som har fler än genomsnittet (låg hållbarhetsprestation) och företag som inte har kontroverser eller färre än genomsnittet (hög hållbarhetsprestation). En maskininlärningsalgoritm lär sig att identifiera mönster hos ca 100 hållbarhetsindikatorer från respektive område E, S och G som är typiska för hög hållbarhetsprestation respektive typiska för låg hållbarhetsprestation. Maskininlärningsalgoritmen konstruerar en separat modell av sambanden mellan dessa totalt ca 100 hållbarhetsindikatorerna per område och förekomst av hållbarhetskontroverser för respektive E, S och G, dvs tre separata betygsmodeller. I vår modellering är det sambanden mellan kontroverser och hållbarhetsindikatorer som styr hur stor vikt som tillmäts enskilda indikatorer, varför vårt sätt att beräkna ESG-betyg är oberoende av subjektiva uppfattningar av hur viktiga enskilda företeelser är. Modellen är därför inte beroende av tillgång till forskning om finansiell materialitet eller subjektiva gissningar. I denna betydelse är modellens vikter av hållbarhetsindikatorer objektiv.

Tidskriftsartiklar i sammanfattning

Totalt är forskningsleveransen sju artiklar för review-tidskrifter, två bokkapitel och ett redaktörskap för en handbok med 60 kapitel.

En första studie, Jan Svanberg, Tohid Ardeshiri, Isak Samsten, Peter Öhman, Tarek Rana & Mats Danielson *Prediction of environmental controversies and development of a corporate environmental performance rating methodology* beskriver en modell för prediktion av miljökontroverser. Artikeln, skickad till *Earth System Governance*, diskuterar metodologiska svagheter med traditionella miljöbetyg och föreslår en ny metod för miljöbetyg som använder maskininlärning för att uppskatta sannolikheten att företag uppfyller miljöförpliktelser och för att förutsäga miljökontroverser. Studien finner att kontroverser kan förutsägas med hög träffsäkerhet och att vår föreslagna metod för beräkning av miljöbetyg överensstämmer med institutionella investerares informationsbehov eftersom de beräknade betygen identifierar företag som är kontroversbenägna. Vi finner också att icke-linjära modeller är mer lämpliga än linjära modeller att identifiera företag som är benägna till kontroverser.

En andra studie, Jan Svanberg, Tohid Ardeshiri, Isak Samsten, Presha Neidermeyer, Peter Öhman, Frank Maisano, Mats Danielson *Supervised Learning for Developing a Social Performance Rating*, skickad till *Decision Science*, beskriver en modell för prediktion av sociala kontroverser. Studien undersöker utformningen av en ny metod för att bedöma den sociala komponenten i ESG-betyg genom att använda maskininlärning som åstadkommer en betygmodell förankrad i sociala kontroverser. Studien visar hur en ny metod kan förbättra beräkningen av sociala hållbarhetsbetyg och belyser orsakerna till den bristande prestanda hos traditionella sociala hållbarhetsbetyg som dokumenterats i tidigare forskning. Den föreslagna maskininlärningsbaserade metoden har bättre förutsäggelseförmåga, vilket ökar den potentiella möjligheten för socialt ansvarsfulla investeringar. Resultaten kommer sannolikt att påverka hur institutionella investerare använder artificiell intelligens för att uppnå högre precision, större noggrannhet i hanteringen av ESG-data och därför också deras portföljer.

En tredje studie Jan Svanberg, Tohid Ardeshiri, Isak Samsten, Peter Öhman, Presha Neidermeyer, Tarek Rana, Mats Danielson & Natalia Semenova *Corporate Governance Performance Ratings with Machine Learning*, skickad till *Intelligent Systems in Accounting, Finance and Management*, beskriver en vår modell för prediktion av bolagsstyrningskontroverser. Studien använder artificiell intelligens (AI) för att föreslå en ny ESG-betygsmetod. Vår metod mäter i vilken utsträckning företag uppfyller bolagsstyrningsförpliktelser med en modell som predikterar bolagsstyrningskontroverser. Vi finner att bolagsstyrningskontroverser kan förutsägas med hög precision, vilket indikerar att bolagsstyrningsindikatorer innehåller information som är relevant för vår betygsmetod och att bolagsstyrningsbetyg med prediktiv validitet kan utvecklas med den föreslagna metoden. Vidare identifierar vi orsakerna varför bolagsstyrningsbetyg baserade på icke-linjära maskininlärningsmodeller kan vara mer tolkbara även än de linjära modeller som används i traditionella ESG-betyg.

En fjärde studie Jan Svanberg, Tohid Ardeshiri, Isak Samsten, Peter Öhman, Presha Neidermeyer, Tarek Rana and Mats Danielson *Objective, Interpretable Environmental Performance Ratings Using Machine Learning*, skickad till *Journal of Sustainable Finance & Investment, special issue, Artificial*

Intelligence and Sustainable Finance, diskuterar varför traditionella miljöbetyg har låg validitet, låg noggrannhet och är svårtolkade för ESG-analytiker. Studien undersöker om en ny metod för miljöbetyg löser kärnproblem med sådana betyg som i slutändan begränsar deras tolkningsbarhet. Vår metod löser tre tolkningsproblem: (1) Traditionella miljöbetyg använder linjära modeller trots att data troligen är icke-linjära, vilket orsakar felaktiga modeller med låg tolkningsbarhet. (2) Traditionella betygssättningsmodeller är lika- eller godtyckligt viktade modeller som saknar begreppslig grund. Avsaknaden av begreppslig grund gör traditionella miljöbetyg godtyckliga och därför mycket svårtolkade. (3) Traditionella miljöbetygsmodellens storlek gör dem i praktiken omöjliga att tolka. Studien föreslår en maskininlärningsbaserad miljöbetygsmetod som maximerar betygets tolkningsbarhet genom att vara begreppsligt grundande och genom att använda SHAP-diagram som möjliggör en lättillgänglig tolkning av miljöbetyg.

I en femte studie Tarek Rana, Md. Hafij Ullah, Faruk Bhuiyan, Parvez Mia, Jan Svanberg & Peter Öhman *Textual analytics and machine learning in examining and predicting sustainability disclosure by local government authorities: Evidence from Australia*, skickad till ***Qualitative Research in Accounting & Management, special issue, On Humans Algorithms and Performance***, undersöks hur lokala myndigheters hållbarhetsrapporter kan analyseras automatiskt med textanalys och maskininlärningsanalys och om framtida förutsägelser av myndigheternas hållbarhetsprestation kan göras. Med hjälp av maskininläringsteori genomförs textanalys och maskininläring för att kartlägga information om hållbarhet från totalt 231 årsredovisningar från 77 lokala myndigheter i delstaten Victoria. Studien finner att textanalys är lämplig för att extrahera och visualisera hållbarhetsredovisning, upptäcka nya ämnen relaterade till hållbarhet och utveckla abstrakta ämnen från ett stort antal dokument. Resultaten belyser vidare potentialen för maskininlärningsstrategi för att utveckla kollektiv kunskap om hållbarhetsmetoder för lokala myndigheter.

I en sjätte studie diskuterar Ulrik Franke & Jan Svanberg, *Cyber incident disclosure, lessons from CO2 disclosure*, skickad till ***Journal of Cyber Policy***, vilka effekter som kan uppstå för företag som råkar ut för cyberincidenter. Cyberincidenter inträffar hela tiden, orsakade av dålig design eller dåliga incitament. För att ta itu med den senare orsaken, har avslöjande av cyberincidenter har föreslagits. Precis som vid hållbarhetskontroverser uttrycker cyberincidenter att det finns brister hos företaget som är inblandat vilket betyder att incidenten är en signal, snarare än det faktiska problemet. Genom bättre redovisning av strukturella svagheter och styrkor skulle cyberincidenter kunna förebyggas men det kan ske till en kostnad för de företag som redovisar sina svagheter. Mot bakgrund av erfarenheterna av koldioxidutsläppsredovisning finner studien att ökad redovisning av cyberincidenter eller de omständigheter som orsakar dem skulle kunna leda till ökade kostnader för eget kapital och skuld för företag med många cyberidentiteter eller med allvarliga cyberincidenter, aktieägaraktivism och eventuellt minskande efterfrågan. Dessa effekter kommer att vara mindre för avslöjande av cyberincidenter än motsvarande effekter för offentliggörande av koldioxidutsläpp.

I en sjunde artikel Jonathan Fluharty-Jaidee, Presha Neidermeyer, Davide Cervone & Jan Svanberg *The Readability of Financial Statements, Information Asymmetries and Managerial Compensation*, skickad till ***The British Accounting Review***, undersöker författarna hur

informationsasymmetri och kompensation till ledningen påverkar utformningen av finansiella rapporter. Studiens underlag är amerikansk redovisning som analyseras med statistiska metoder.

Böcker och bokkapitel

Tarek Rana, Alan Lowe & Jan Svanberg är huvudredaktörer för Springer *Handbook of Big Data and Analytics in Accounting and Auditing*, om 60 kapitel med bidrag från hela världen. Handboken är det första stora samlingsverket om användningen av artificiell intelligens i redovisning och revision. Jan Svanberg, Peter Öhman och Presha Neidermeyer är dessutom redaktörer för sektionen om ESG-betyg bestående av sex kapitel med bidrag bl a från Fredrik Nilsson, Professor Uppsala universitet och ordförande i Svenska Revisionsakademien samt med i vår referensgrupp för P8/18, och med bidrag av Beatrice Crona, från det världsledande *Stockholm Resilience Center*. Beatrice har ett Vinnova-projekt om varför biosfären inte inkluderas i ESG-betyg, är mycket intresserad av ESG-betyg och vi räknar Beatrice till vårt nätverk.

Jan Svanberg, Tohid Ardeshiri, Isak Samsten, Peter Öhman, Presha Neidermeyer & Mats Danielson har i handboken författat kapitlet *High Precision ESG ratings with Artificial Intelligence*. Kapitlet sammanfattar resultaten från E, S, och G och diskuterar hur institutionella investerare kan förbättra sina hållbarhetsinvesteringar genom att använda ESG-betyg som kan identifiera vilka företag som löper hög risk att bli föremål för hållbarhetskontroverser.

Slutligen: Alan Montague, Jan Svanberg, Frank Maisano, & Venesser Fernandes är författare till Chapter 6: Financial and Insurance Services, i John Burgess, Julia Connell, Alan Montague, Alan Nankervis, *The Fourth Industrial Revolution: Impacts on Australian Industry & Future Challenges*, Springer, 2021. Kapitlet behandlar utmaningar och möjligheter för finansindustrin i Australien med användandet av artificiell intelligens. Författarna hävdar att artificiell intelligens kan bli en konkurrensfördel för den mycket stora australiensiska finansindustrin men diskuterar också hur sysselsättningen kan påverkas negativt genom rationalisering av administrativ produktion. Det är en ära att få vara med i boken som riktar sig till den australiensiska regeringen.

Praktisk Leverans

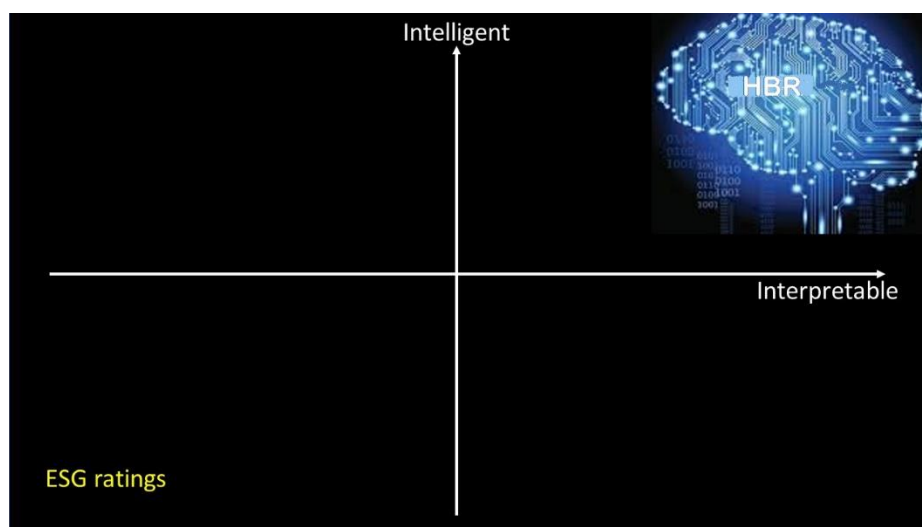
HBR med Prediktiv Modellering

Isak Samsten, Jan Svanberg men senare huvudsakligen Tohid Ardeshiri, har utvecklat en maskininlärningsbaserad metod att beräkna ESG-betyg som vi kallar Hållbarhetsroboten (HBR). Vi beskriver hur den ser ut och hur den fungerar. HBR ger tre separata betyg för alla företag, i nuvarande utförande 2517 företag men i den version som vi i egen regi kommer att utveckla under tidig höst finns

upp emot 10 000 företag analyserade. Utseende och funktion är ungefär detsamma för de E, S, och G så vi beskriver enbart E.

Målet med HBR är att använda artificiell intelligens för att konstruera en metod som är bättre än traditionella betygsmetoder på två punkter: (1) Intelligens, vilket för ett ESG-betyg är förmågan att skilja mellan hållbara och ohållbara företag med hög träffsäkerhet. (2) Tolkningsbarhet, vilket för ett ESG-betyg är förmågan att förklara varför ett företag är hållbart eller ohållbart. Figur 1 illustrerar hur vi uppfattar dels traditionella ESG-betyg (i bilden betecknade ESG ratings) enligt den forskning vi tagit del av, dels önskad positionering av HBR. Traditionella betyg är beräknade med linjära, manuellt utvecklade modeller med ett relativt fåtal indikatorer för vilka validitet är starkt ifrågasatt av tidigare forskning. Vi anser att dessa modeller som gör anspråk på att mäta så komplexa företeelser som företags hållbarhet inom miljö, samhälle och bolagsstyrning inte är ändamålsenligt utformade.

Traditionella betygsmodeller är typiska exempel på *explanatory* (förklarande) modellering. Syftet med förklarande modellering är att förklara samband mellan begrepp som gäller för flertalet företag i en population. Denna typ av modellering är därför lämpad för teoriutveckling. Ett ESG-betyg är emellertid ingen teori, utan diagnostik av enskilda, specifika företag. Denna uppgift är lämpad enbart för prediktiv modellering, vars syfte är att fälla *träffsäkra omdömen om enskilda fall* (Shmueli, 2010). En förklarande modell anses bra om den teoretiskt förankrat beskriver ett genomsnittsföretag medan en prediktiv modell bedöms enbart på sin förmåga att identifiera, i vårt fall, hållbara företag med träffsäkerhet. Av detta skäl använder vi, till skillnad från betygsindustrin, prediktiv modellering med maskininläring som metod att konstruera ESG-betyg. Prediktiv modellering har högre kapacitet att representera komplexitet varför vi positionerar HBR högt på y-axeln och av skäl som förklaras under rubriken 'Tolkning av betyg' är HBR också långt åt höger på x-axeln medan traditionella modeller ligger långt åt vänster.



Figur 1. Positionering av HBR i förhållande till traditionella ESG-betygsmodeller.

Precision och prediktiv validitet

Vi har nämnt under forskningsleverans att metoden bakom HBR skiljer sig från traditionella metoder på den viktigaste tänkbara punkt: Den har prediktiv validitet. Vi illustrerar med Tabell 1 som härstammar från vårt bokkapitel i Handbook of Big Data and Analytics in Accounting and Auditing. I tabellen visas fem vanligt förekommande prestationsmått som mäter prediktionsförmåga hos maskininlärningsalgoritmer. Vi betraktar endast den första – precision. Precision betyder förmågan hos vår Random Forest maskininlärningsmodell att korrekt identifiera de företag som har minst en kontrovers (för miljö och 'Governance' = bolagsstyrning) eller som har fler kontroverser än genomsnittsföretaget (för social eftersom det där finns så många kontroverser) i förhållande till antalet företag som pekas ut som kontroversföretag av modellen. Om Random Forest alltså kastar 10 pilar på en tavla och träffar sju gånger är precisionen 0,7. Vi ser att både den linjära modellen och HBRs Random Forest kan identifiera kontroversföretag med hög precision jämfört med vad slumpmässigt gissande skulle resultera i, som för miljö skulle innebära en precision på ca 0,12 eftersom ca 12 procent av företagen haft en miljökontrovers under den tioårsperiod vi har som bedömningsfönster. Precisionsvärdena på ca 0,7 (70 procent precision) eller högre visar att det finns information i hållbarhetsindikatorerna som vår metod kan använda för att sätta betyg på företagen. De visar också att det är möjligt att med gott resultat identifiera företag som sannolikt har egenskaper typiska för lågt hållbarhetspresterande företag, dvs typiska för företag som kritiserats för bristande hållbarhet. Även om i enskilda fall ett företag skulle kunna vara på falska grunder utpekad för brister i hållbarhet för en genomsnittlig hållbarhetskontrovers är det sannolikt att, sett över 10000 - 20000 kontroverser, dessa kontroverser indikerar brister i hållbarhet. En metod som kan identifiera kontroversbenägna företag med prediktiv validitet bör därför vara en metod att mäta bristande efterlevnad av hållbarhetskrav med validitet.

Tabell 1: Prediktionsförmågan hos Random Forest (som används i HBR) jämfört med logistisk regression som är en vanlig linjär modell.

Environment	Precision	Recall	F-measure	AUC	PRC
Logistic Regression	0,3045	0,7253	0,4282	0,8727	0,4814
Random Forest	0,6994	0,1942	0,2990	0,8849	0,5090
Social					
Logistic Regression	0,6132	0,7788	0,6852	0,8982	0,7898
Random Forest	0,8417	0,5756	0,6823	0,9165	0,8104
Governance					
Logistic Regression	0,3387	0,6513	0,4444	0,7531	0,4699
Random Forest	0,7506	0,1700	0,2756	0,7787	0,5110

Betygsexempel och beskrivning av funktion

Företagen som analyseras av HBR är från hela världen, men huvudsakligen från Europa och USA. I Tabell 2 ser vi de 13 företag som får absolut sämst miljöbetyg av HBR.

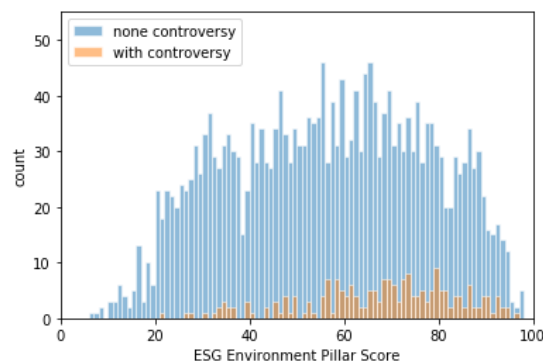
Tabell 2: Exempel på de miljömässigt sämsta företagen

Företag	Land	Miljöbetyg	Konservativt miljöbetyg	E-score Refinitiv	# Miljökontrov	# Sociala kontroverser	# Bolagsstyrning	Market Cap [\$B]	Pressfrihetsindex	GICS Sektor
Enel SpA	IT	0	0	83	3	6	0	67.2	76	Utilities
QEP Resources Inc	US	0	0	42	0	1	0	0.9	76	Energy
Electricite de France SA	FR	0	0	63	0	28	0	32.6	77	Utilities
First Quantum Minerals Ltd	CA	0	0	63	1	3	1	5.3	85	Materials
OMV AG	AT	0	0	74	0	9	2	15.9	84	Energy
Mechel PAO	RU	0	0	47	0	3	0	0.6	51	Materials
Teck Resources Ltd	CA	0	0	69	1	1	0	9	85	Materials
Rio Tinto Ltd	AU	0	0	84	5	25	3	81.8	80	Materials
Vale SA	BR	0	0	78	4	17	1	56.2	66	Materials
Petroleo Brasileiro SA Petrobras	BR	0	0	60	5	30	9	83.1	66	Energy

Från vänster står företagsnamn, land, miljöbetyg enligt vår modell, konservativt miljöbetyg enligt vår modell, E-score Refinitiv är ett miljöbetyg från mediaföretaget Refinitiv, # miljökontroverser företaget har haft under en period av 10 år, # sociala kontroverser och bolagsstyrningskontroverser, marknadsvärde, pressfrihetsindex, och sektor enligt systemet GICS. HBR ger två miljöbetyg. Det ena är det som genereras automatiskt från Random Forest-modellen, som är ett tal mellan 0 och 100 där 0 är mycket hög sannolikhet för att företaget inte tar sitt miljöansvar, vilket betyder att företaget kan förväntas bryta mot miljöregler och 100 är en låg sannolikhet att bolaget bryter mot miljökrav (rättsliga eller moraliska). Alla de 13 sämsta företagen har betyg 0. Eftersom vår modell har lärt sig att känna igen hur företag typiskt ser ut som har kontroverser och eftersom kontroverser förutsätter att det finns en effektiv mediabevakning är sannolikheten att ett företag blir granskat av media beroende av pressfrihet. I länder där pressfriheten är hög fungerar granskningen bra men i länder där pressfrihet nästintill inte förekommer alls såsom i Kina, som har världens lägsta pressfrihet enligt World Press Freedom Index utgivet av organisationen Journalister utan gränser, fungerar granskningen av företag sämre. I länder med låg pressfrihet har kontroverserna ett mörkertal som är beroende av pressfrihet och företagsstorlek. Stora företag har normalt högre mediabevakning än små företag varför effekten av låg pressfrihet blir särskilt stor för de största företagen. Vi har därför konstruerat ett miljöbetyg som har en säkerhetsmarginal för skillnader mellan länder i pressfrihet. Det konservativa miljöbetyget är vårt ursprungliga miljöbetyg reducerat för en kombinerad pressfrihets- och företagsstorleksfaktor som är olika stor för olika företag. Stora företag i Kina har störst skillnad mellan miljöbetyget i kolumn tre och fyra (vilket inte syns hos de 13 miljömässigt sämsta företagen eftersom de alla har noll i båda kolumnerna). Det konservativa miljöbetyget är alltså det som lämpar sig bäst vid jämförelse av företag i olika länder.

Varför har E-score Refinitiv tämligen goda betyg på dessa företag medan HBR ger dem en nolla? Förklaringen är de problem vi identifierat med traditionella ESG-betyg. Refinitiv använder ett lika-viktat medelvärde av ett stort antal miljöindikatorer, vilket betyder att viktiga omständigheter

värderas på samma sätt som de indikatorer som rätteligen borde vara betydelselösa. Totalt använder Refinitiv 61 miljöindikatorer från tre områden, resursanvändning, utsläpp och innovation som mått på miljöprestation. En betygssättningsmodell som tar ett aritmetiskt medelvärde av så många indikatorer tenderar att förlora den mesta av informationen hos indikatorerna när 61 mätetal ersätts av ett enda. Vi har funnit att Refinitivs miljöbetyg inte alls avspeglar informationen som finns i miljökontroverser. Vi illustrerar detta i Figur 2. Eftersom fördelningarna i Figur 2 är tämligen lika för företag med respektive utan miljökontrovers har Refinitivs miljöbetyg ingen information företagsegenskaper typiska för kontroversbenägna företag. Problemet med en modell som bedömer miljöprestation utan relation till miljökrav kan liknas vid en trafikskola som bedömer en förarens förmåga att framföra sitt fordon enbart efter hur fort eller långsamt föraren kör, oavsett hastighetsbegränsning.



Figur 2: Fördelningar av företag med miljökontrovers (gula staplar) och utan kontrovers (blå staplar). X-axeln har miljöbetyget från Refinitiv och y-axeln är antalet företag.

Rio Tinto är tex ett företag som dragit till sig negativ uppmärksamhet när det gäller respekt för både miljöregler och ursprungsbefolkningars rättigheter och driver gruvor i sektorn 'Materials'. Av Refinitiv får Rio Tinto 84 av 100 eftersom företaget genom sin storlek åtminstone på papperet har ett stort antal policies och liknande åtgärder som ger intryck av ett företag som söker minska sina miljörisiker. HBR ger Rio Tinto noll. HBR har förmågan att kritiskt bedöma betydelsen av policies och skenbart riskreducerande åtgärder i ljuset av vad som händer i liknande företag (Se avsnittet Tolkning av betyg). Tex ter sig indikatorn "Biodiversity Impact Reduction" som positiv för Refinitivs modell eftersom den tycks kunna signalera att företag har gjort något för att minska sin påverkan på växter, insekter eller djur men HBR tolkar ett högt värde på "Biodiversity Impact Reduction" som en indikation på att företaget sannolikt brister i sitt miljöansvar eftersom företag som har höga värden på denna indikator ofta har miljökontroverser.

I Tabell 3 visas 14 företag som hamnat i mitten av betygsskalan. Noterbart för dessa företag liksom för de som hamnat på noll är att betyget inte är detsamma som att ha haft miljökontroverser. Ett företag som inte haft kontroverser, som är fallet QEP Resources Inc i USA, kan ha mycket lågt betyg trots att det inte självt varit inblandat i överträdelser av miljökrav. Det beror på att det i flera avseenden liknar företag som har haft en eller flera miljökontroverser. Ju högre miljöbetyg desto mindre vanligt är det emellertid att finna miljökontroverser eftersom det är sannolikheten för miljökontroverser som är grund

för miljöbetyget. (Notera dock att detta inte alls stämmer för Refinitivs miljöbetyg i Figur 2). Vi finner Telia AB på 66 trots att Telia inte haft någon miljökontrovers.

I Tabell 3 kan vi också se hur justeringen för pressfrihet fungerar. Japan har ett något lägre pressfrihetsindex än Sverige, 71 jämfört med 91, vilket gör att japanska företag har en större skillnad mellan miljöbetyget och det konservativa miljöbetyget än svenska företag har. Det beror på att de japanska företagen utsätts för mindre närgången granskning av miljöaspekten av sin verksamhet än de svenska företagen, varför det konservativa miljöbetyget måste ta höjd för mörkertalet. Det är mer sannolikt att japanska företag har brister i miljöansvar utan att upptäckas av media än det är att svenska företag har det. Eftersom det framför allt är de allra största företagen som granskas av media behöver justeringen för pressfrihet vara extra stor om företaget är mycket stort. Nidec Corp har ett miljöbetyg på 66 men ett konservativt miljöbetyg på 57. I Telias fall är motsvarande skillnad endast 66 mot 65. Sverige ligger tillsammans med övriga nordiska länder den internationella toppen i pressfrihetsindex, varför miljöbetygen har en lägre osäkerhet här än för företag från länder med låg pressfrihet såsom Kina med 22 i pressfrihet. I många fall går pressfrihet hand i hand med ekonomisk utveckling, tex Egypten som har 43 i pressfrihet, men även ekonomiskt utvecklade länder som Singapore (45) kan ha låg pressfrihet. Vid jämförelse av miljöprestationen för företag i tex norra Europa och USA med företag i Asien och Afrika är det därför relevant att använda det konservativa miljöbetyget.

Tabell 3: Exempel på företag med ett miljöbetyg strax över skalans mitt.

Företag	Land	Miljöbetyg	Konservativt miljöbetyg	E-score	#	# Sociala	#	Market Cap [\$B]	Pressfrihetsindex	GICS Sektor
				Refinitiv	Miljökontrov	kontroverser	Bolagsstyrning			
Ryanair Holdings PLC	IE	65	64	37	0	31	1	11.2	87	Industrials
Bolsas y Mercados Espanoles SHMSF	ES	65	62	65	0	0	0	1.8	78	Financials
Telia Company AB	SE	66	65	72	0	13	0	17.4	91	Communication Services
Gerdau SA	BR	66	57	60	0	4	1	4.8	66	Materials
Iron Mountain Inc	US	66	61	57	0	6	0	8.3	76	Real Estate
Nidec Corp	JP	66	57	54	0	0	0	37.9	71	Industrials
NGK Spark Plug Co Ltd	JP	66	60	69	0	3	0	3.7	71	Consumer Discretionary
Bank of Montreal	CA	66	63	88	0	12	0	42.4	85	Financials
Acer Inc	TW	66	62	71	0	5	1	1.6	76	Information Technology
DXC Technology Co	US	66	61	77	0	3	4	7.6	76	Information Technology
Caesars Entertainment Corp	US	66	61	65	0	0	0	7.3	76	Consumer Discretionary
Mitsubishi Estate Co Ltd	JP	66	57	79	0	0	0	24.5	71	Real Estate
BWP Trust	AU	66	63	64	0	0	0	1.5	80	Real Estate
Yum! Brands Inc	US	66	60	56	0	36	2	31.1	76	Consumer Discretionary

Eftersom samhällets syn på företag av uppenbara skäl skiljer sig åt mellan sektorer är det rimligt att miljöbetygen också skiljer sig åt markant mellan sektorer. Dagens miljödebatt är koncentrerad till klimatet varför sektorer som har höga utsläpp av koldioxid av samhället anses mer miljöskadliga än andra sektorer. Många institutionella investerare som LF Kapital använder därför sektorskillnader som ett första kriterium vid bedömning av investeringars hållbarhet. Ett relativt miljöbetyg som Refinitiv E-score bedömer företagen, sektor för sektor, i relation enbart till andra företag i samma sektor vilket leder till att betygen inte kan användas för att jämföra sektorer eller företag i olika sektorer. HBRs miljöbetyg är ett

absolut omdöme eftersom den beräknar sannolikheten för brott mot miljökrav för varje företag, och denna sannolikhet är oberoende av vad som försiggår i andra företag. HBR bedömer därför sannolikheten att ett visst företag lever upp till samhällets krav på miljöhänsyn. HBRs bedömning är baserad enbart på fakta, och fakta talar ett mycket tydligt språk, vilket framgår av sektorjämförelsen i Tabell 4.

Tabell 4: Sektorjämförelse

GICS Sector Name	Proposed env rating	conservative env rating	E-score Referitiv	# env controversies	# soc controversies	# gov controversies	Market Cap [\$B]
Energy	11	8	76	130	881	73	3140
Utilities	13	9	64	59	454	14	1748
Materials	29	25	67	96	892	54	2488
Consumer Staples	31	26	72	16	1100	48	4321
Information Technology	42	35	75	2	1118	103	6740
Communication Services	44	31	69	1	1213	74	4192
Health Care	46	41	70	7	1209	110	4572
Consumer Discretionary	51	44	67	36	1896	103	4984
Industrials	51	44	67	19	1746	99	5048
Financials	57	43	67	5	2369	310	8986
Real Estate	84	75	60	1	130	16	1453

I sektoruppdelningen som används av LF Kapital och många andra, GICS, finns 11 sektorer. Många har självförklarande namn men andra är inte helt självklara: 'Consumer staples' är dagligvaror som mat och dryck, 'Consumer Discretionaries' är sällansköpsvaror såsom diskmaskiner och bilar. 'Utilities' är leverans av energi, naturgas och vatten, medan 'Energy' är de företag som utforskar, producerar, förädlar, marknadsför, lagrar och transporterar olja och gas, kol och andra bränslen. Det framgår av tabellen att 'Energy' och 'Utilities' är sektorer med markant högre risk för brott mot miljökrav jämfört med de andra sektorerna men 'Materials' och 'Consumer staples' har också tämligen höga miljörisker. Av uppenbara skäl finns låga miljörisker i 'Real Estate' och 'Financials'.

Sektor är dock endast den översta nivån av branschuppdelning. Motsvarande översikter och jämförelser kan göras för mer finkorniga uppdelningar mellan industrigrupper, industrier och underindustrier (av vilka det finns så många som 157). Jämförelsen mellan underindustrier avslöjar att det finns grupper av verksamheter som har mycket hög risk för bristande miljöansvar. Tabell 5 återger 25 underindustrier med högst miljörisk. När kommande produkter kan betygssätta den större gruppen av 10 000 företag kan denna typ av jämförande översikter vara ett effektivt verktyg för ESG-analys av portföljer. Kommande versioner av produkter baserade på kontroversprediktionsmetoden att bedöma hållbarhet kommer dock att ha inbyggda funktioner för analys av givna portföljer. Antag att en portfölj hos LF Kapital består av 1500 företag. HBR skulle kunna läsa in portföljen och visa i grafer var potentialen för förbättrad hållbarhetsprestation är störst, dvs var portföljen är mest exponerad för hållbarhetsrisker.

Tabell 5: Översikt över underindustrier med lägst miljöbetyg.

GICS Sub-Industry Name	Proposed env rating	conservative env rating	E-score Refenitiv	# env controversies	# soc controversies	# gov controversies	Market Cap [\$B]
Integrated Oil & Gas	2	1	82	57	469	30	1894
Diversified Metals & Mining	3	3	76	32	200	13	495
Renewable Electricity	8	3	70	0	0	0	17
Electric Utilities	9	6	64	27	249	6	1022
Personal Products	11	7	79	3	105	7	583
Automobile Manufacturers	11	8	82	23	567	33	702
Multi-Utilities	12	8	68	21	150	5	436
Independent Power Producers	12	6	49	5	21	2	70
Coal & Consumable Fuels	13	9	77	14	40	2	76
Systems Software	15	9	87	0	123	10	1230
Forest Products	17	15	61	0	4	0	12
Gold	17	15	58	18	86	13	183
Hypermarkets & Super Center	17	14	91	0	146	4	490
Copper	18	15	68	8	31	7	66
Aluminum	18	13	62	2	8	0	18
Oil & Gas Refining & Marketing	19	15	74	5	38	8	245
Oil & Gas Exploration & Production	20	13	60	35	156	19	514
Precious Metals & Minerals	20	12	73	2	26	0	30
Brewers	21	18	61	0	49	1	291
Industrial Gases	21	18	82	0	11	0	201
Technology Hardware, Storage & Peripherals	21	14	76	1	368	24	1332
Food Distributors	23	16	73	0	14	0	42
Steel	24	20	72	15	236	10	296
Fertilizers & Agricultural Chemicals	24	17	42	5	29	3	64

Tolkning av betyg

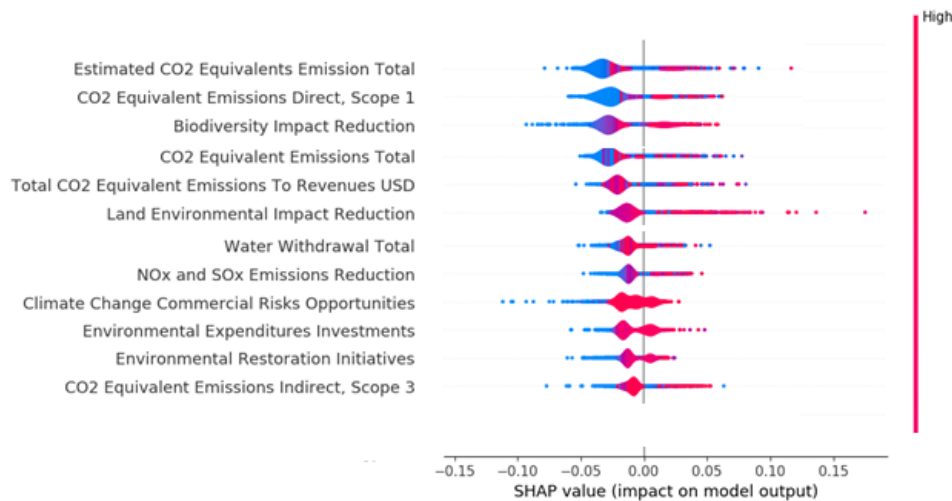
HBR presenterar i nuvarande utformning betyg, men vi har undersökt förutsättningarna för att använda grafiska presentationer av företag som gör det möjligt att snabbt förstå orsakerna till att robotens bedömningar. I kommande produktutveckling kommer varje företag och varje betyg att kopplas till en databas med grafer så att en komplex analys av ett företags hållbarhet kan ske på ett överskådligt sätt och med omedelbarhet. Figur 3 visar ett exempel på hur grafisk presentation förklarar skälen till HBRs hållbarhetsbedömning. Figuren kallas ”SHAP Summary Plot” och används tillsammans med metoder, som enligt vår uppfattning är de mest användbara för vår tillämpning, utvecklade av Scott Lundberg, se <https://scottlundberg.com/> (Senior Researcher at Microsoft Research och Affiliate Assistant Professor at the University of Washington) och hans kollegor (Lundberg et al., 2019; Lundberg and Lee, 2017). Lundberg och kollegor är pionjärer på metodutveckling för tolkningsbar artificiell intelligens med tillämpning på problem inom medicin och hälso- och sjukvård. Detta har lett till utveckling av allmänt tillämpliga metoder och verktyg för att tolka komplexa maskininlärningsmodeller som nu används inom bank, logistik, tillverkning, molntjänster, ekonomi, sport och andra områden. Vi använder dessa för att förklara vår Random Forest-modell som med hjälp av Lundbergs SHAP blir mer tolkningsbar än någon annan tillgänglig modell, trots att modellen samtidigt tillhör de mest komplexa modellerna.

Figur 3 visar att för ett genomsnittligt företag är det koldioxid, biodiversitet, återställande av skador på mark, och vattenanvändning som är viktiga omständigheter. Roboten fäster dock olika vikt vid

olika faktorer vid bedömning av olika sektorer, industrigrupper, industrier, underindustrier, ända ner till enskilda företag och grafen i Figur 3 visar endast genomsnittsföretaget. Varje företag granskas genom tusentals jämförelser med andra företag på ett sätt som i varje enskilt fall leder till ett unikt betygsbeslut och ett unikt SHAP-diagram. Varje företag har alltså ett unikt diagram för varje betyg E, S och G. Till skillnad från en statisk modell som använder samma betydelsevikter för alla företag i tex en sektor, fungerar HBR ungefär på samma sätt som hur en domare utmäter ett straff med beaktande av omständigheter i ett enskilt fall. Mindre komplexa, statistiska modeller kan inte uppnå samma precision i identifikationen av hållbarhetsrisker eftersom de inte har dynamiken hos bedömningssättet hos prediktiv modellering med komplexa maskininlärningsalgoritmer.

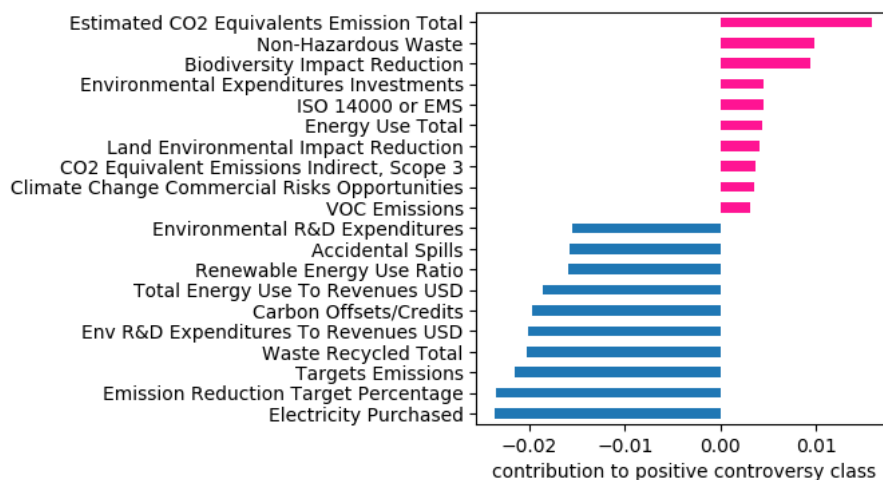
Figur 3 illustrerar översiktligt hur modellens inre logik ser ut när den bedömer företag med diverse miljöegenskaper. I Figur 3 betyder blå färg lågt värde på indikatorn, dvs om det gäller koldioxiden betyder blått värde lågt utsläpp. Den röda stapeln till höger betyder att ju högre upp man kommer desto större betydelse har indikatorerna för miljöbetyget. 'Estimated CO2 Equivalent's Emission Total' har alltså i genomsnitt för alla företag den enskilt största betydelsen för miljöbetyget, varför HBRs uppfattning går i linje med dagens miljödebatt. Varje företag representeras av en liten röd eller blå punkt för varje indikator (dvs varje rad). Det bitvis tjockare strecket på den översta raden består av 2517 punkter där de flesta företagen har en blå punkt sammanklumpade vid SHAP-värde ungefär -0,05. Skalan på x-axeln är storleken på betydelsen av en blå eller röd punkt för ett visst företags klassificering som miljö(o)ansvarigt. Ett exempel på ett företag är den ensamma röda punkten till höger. Den röda färgen betyder högt koldioxidutsläpp. SHAP-värdet på ca + 0,10 betyder att företagets koldioxidutsläpp har varit starkt bidragande till att företaget bedöms ha en hög miljörisk, till skillnad från de många blå punkter som klumpar sig till vänster, indikerande att låga koldioxidutsläpp istället är associerade med hög sannolikhet för efterlevnad av miljökrav, dvs låg miljörisk.

En ”SHAP Summary Plot” kan visa många omständigheter i samma bild och är ett sätt att synliggöra hur HBR bedömer en grupp av företag, tex alla företag i databasen, alla företag i en sektor, eller del av sektor, eller i en portfölj. På x-axeln finns SHAP-värde som är hur stor betydelse enskilda mätvärden har haft för miljöbetyget. På y-axeln syns indikatorers betydelse för betyget, ju högre upp desto större betydelse. Höga och låga mätvärden visas med rött och blått och antalet företag som har ett indikatorvärde med samma SHAP-värde visas som klumpar av punkter, vilket ger en känsla för fördelningen av företag över SHAP-värden.



Figur 3: SHAP Summary Plot som visar hur HBR klassificerar företag efter deras efterlevnad av miljökrav. Bilden visar modellens bedömning av alla företag där varje företag representeras av en punkt per indikator. Endast de 12 indikatorerna med störst betydelse visas. Figuren visar den relativa betydelsen för dessa 12 indikatorer vid bedömning av alla företag. Den relativa betydelsen för ett enskilt företag är inte densamma som den genomsnittliga betydelsevikten. SHAP Summary Plot kan användas för att förstå en sektor eller del av en sektor.

För att visa hur betygen för enskilda företag har beräknats av HBR finns flera typer av diagram och beräkningarna av nyckeltalen som presenteras i graferna kan också skilja sig åt. Ett enkelt alternativ är stapeldiagrammet i Figur 4 som visar vilken betydelse de viktigaste indikatorerna har haft för beslutet att ge ett företag ett visst betyg. Vi känner igen flera av indikatorerna från Figur 3 och kan konstatera att detta företag verkar vara ett genomsnittligt företag från HBRs perspektiv eftersom de viktiga indikatorerna förekommer i bedömningen av detta företag. Denna typ av diagram kan kopplas till alla betyg i HBR.



Figur 4: Stapeldiagram som visar betydelsen av enskilda miljöindikatorer för miljöbetyget. Rosa staplar drar ner betyget och blå drar upp det.

Vid bedömning av enskilda företag kan helt olika omständigheter framträda som de mest betydelsefulla för betyget, eftersom företagens verksamhet skiljer mycket åt. Denna dynamik hos HBRs inre logik tillsammans med HBRs höga precision i identifikationen av företag med lågt hållbarhetsansvarstagande är de två viktigaste skillnaderna mot traditionella betygsmodeller. Därför är det upplysande att se några diagram från olika typer av företag. Samtidigt som vi gör det använder vi också en annan typ av diagram som kallas *SHAP Force Plot*, och som har blivit populärt som verktyg att tolka AI-modeller i datavetenskap. Vi bedömer att *SHAP Force Plot* är den bästa grafiska illustration att länka till varje betyg på en sida med sammanfattande information om det bedömda företaget.

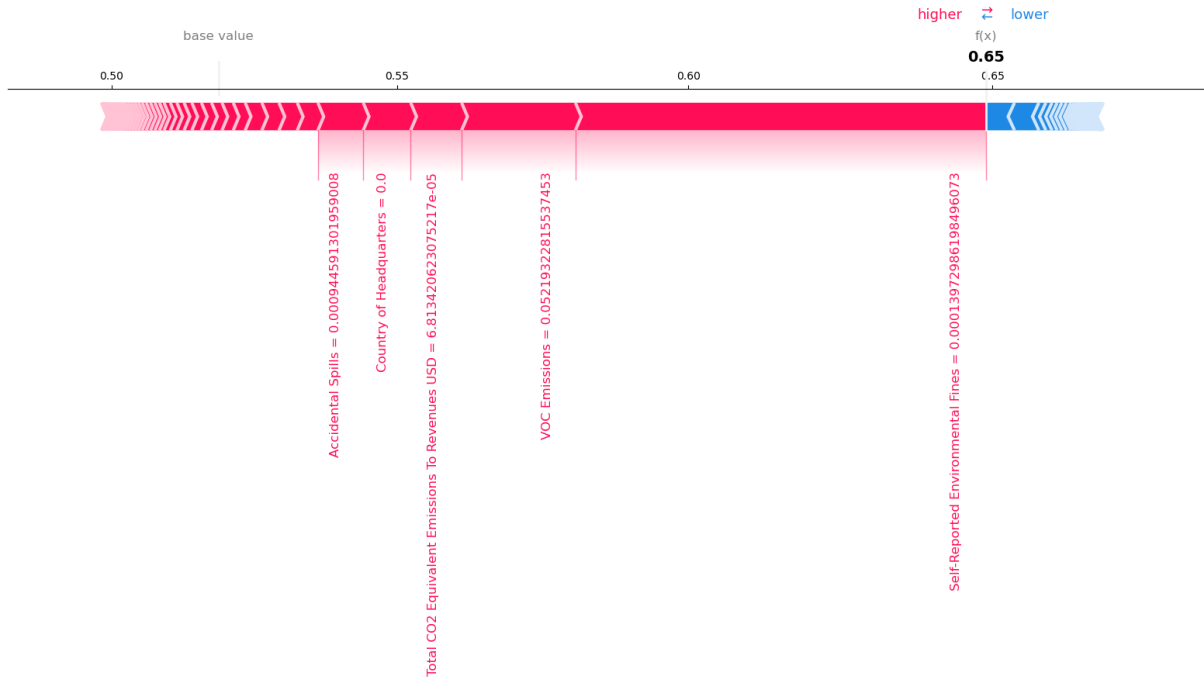
I Figur 5 och 6 och finns två exempel på *SHAP Force Plot*. Vi ser först i Tabell 6 på två tämligen olika företag, Imperial Oil Ltd i Kanada och Taiwan-företaget Delta Electronics Inc som uppenbarligen finns inom IT och som också tycks sätta värde på att profilera sig som miljövänligt företag. Delta har varit med i Dow Jones Sustainability Indices fem år i rad enligt bolagets egen hemsida. HBR har gjort en oberoende granskning av Delta Electronics och funnit att utmärkelsen är relativt välförtjänt även om bolaget inte är helt utan miljörisker. En faktor som gör HBRs bedömning något konservativ är att pressfriheten i Taiwan är 76, vilket är högt men inte i paritet med de allra högsta. Det har lett till att HBR tar höjd för den osäkerheten och ger Delta Electronics 88. Imperial Oil Ltd har däremot ett konservativt miljöbetyg på 3.

Tabell 6: Miljöbetyg för Imperial Oil Ltd i Kanada och Delta Electronics Inc i Taiwan

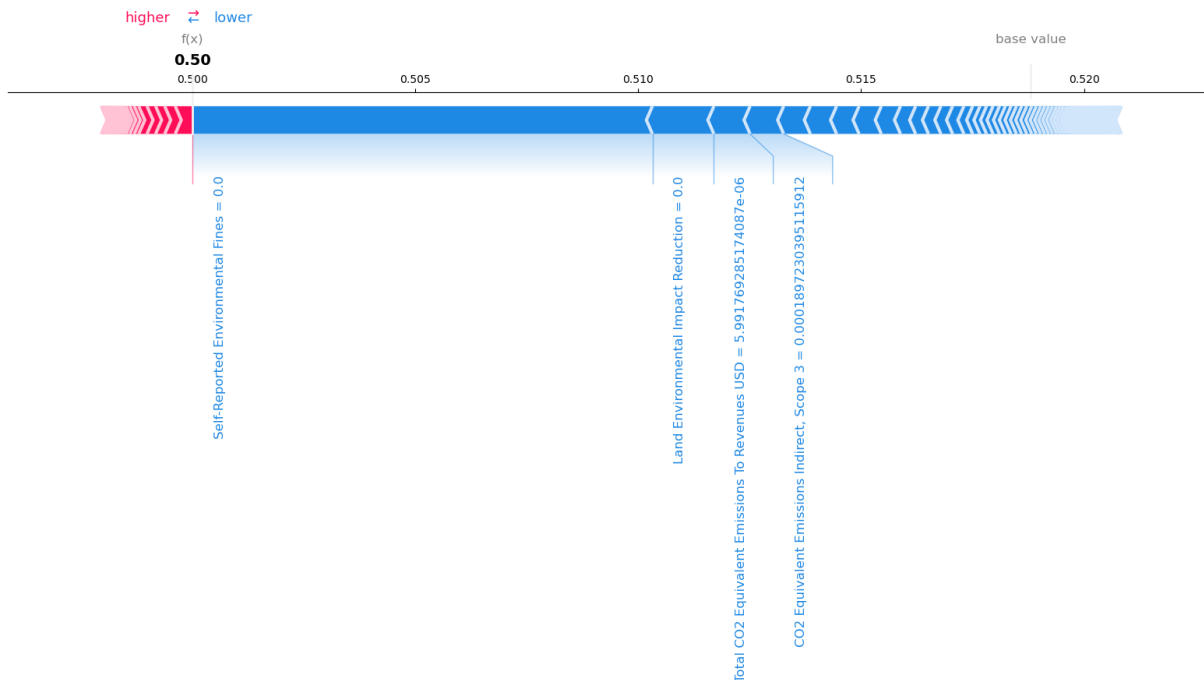
Företag	Land	Miljöbetyg	Konservativt miljöbetyg	E-score Refinitiv	# Miljökontrov	# Sociala kontroverser	# Bolagsstyrning	Market Cap [\$B]	Pressfrihetsindex	GICS Sektor
Imperial Oil Ltd	CA	5	3	79	2	3	0	18.8	85	Energy
Delta Electronics Inc	TW	93	88	82	0	1	0	11.1	76	Information Technology

Figur 5 visar de 'krafter' som drar beslutet att ge det kanadensiska företaget Imperial Oil Ltd ett lågt miljöbetyg. Varje indikatorvärde är en kraft som antingen ökar eller minskar miljöbetyget. Miljöbetyget börjar från baslinjen (mitten). *Bas value* för SHAP-värden är genomsnittet av alla förutsägelser. I diagrammet är varje SHAP-värde en pil som trycker för att öka (positivt värde, blått) eller minska (negativt värde, rött) miljöbetyget. Dessa krafter balanserar varandra vid det faktiska betyget (angivet med en annan skala i figuren än 0-100 som vi använder för betyget). Det röda streckets längd utgör bidraget från alla indikatorer som drar ner Imperials miljöbetyg. De största bidragen är uttryckligen angivna med namn. Som bilden visar är de röda försvårande omständigheterna helt dominerande för miljöbetyget. Det finns mycket små förmildrande omständigheter i den blåmarkerade högra änden av kraftskalan. Vi läser också vilka de försvårande omständigheterna är och noterar att det är utsläpp av koldioxid men också flyktiga organiska föreningar som kallas VOC (och har brukat finnas i färg och lack), och det har förekommit oljespill pga olyckshändelser i verksamheten som kan tyda på kvalitetsbrister. Två andra omständigheter som givit HBR anledning att rikta misstankar mot Imperial är att företaget har drabbats av miljöböter, vilket i sig självt inte behöver vara ett problem men som likt oljespillet kan tyda på att företagets kontrollmekanismer för ansvarstagande för miljön inte är tillräckliga. En sista omständighet som i sig självt till synes inte har något att göra med ansvarstagande för miljö är var företaget har sitt

huvudkontor. Huvudkontorsfrågan måste tolkas av en ESG-expert. Detta exempel illustrerar hur HBR ger ESG-experten hos institutionell förvaltning uppslag till frågor att ytterligare undersöka.



Figur 5: SHAP Force Plot för företaget Imperial Oil Ltd, i sektorn Energy



Figur 6: SHAP Force Plot för Delta Electronics Inc i sektorn Information Technology

Motsvarande *SHAP Force Plot* för Delta Electronics ser ut att vara motsatsen med i huvudsak blå färg. Här är det en blandning av indicier av olika slag som lett till ett högt miljöbetyg. Företaget har inte drabbats av någon miljöbot, har inte heller redovisat några insatser för att minska sin påverkan på landområden, vilket HBR tolkar positivt. Enligt robotens erfarenhet är *Land Environmental Impact Reduction* först och främst ett tecken på att företaget har en miljömässigt riskfylld verksamhet, inte på den goda insatsen företaget gör att återställa markskador. Slutligen är det låga värden på koldioxidutsläpp som bidragit till att ge Delta Electronics ett högt miljöbetyg. Med en högre upplösning av diagrammet går det att se fler indikatorer som påverkat betyget. Det finns ett stort antal mindre bidragande orsaker till betyget. Dessa syns som kortare pilar i diagrammet.

Avslutande kommentar

Hållbarhetsexperter kan med hjälp av de miljöbetyg och de förklarande diagram som HBR kan leverera för varje företag, för varje sektor, industrigrupp, industri och underindustri kombinera HBRs datadrivna analys med expertkunskaper och kompletterande information från andra källor för att göra mer precisa analyser och ställa mer avslöjande frågor än som är möjligt med någon annan analysteknik. Tolkningen av HBRs analyser är ett av flera verktyg som tex LF Kapitals hållbarhetsexperter förfogar över. *HBR bidrar med information som gör det möjligt att ställa bättre frågor snarare än definitiva svar.* Tolkningen av datapunkter för enskilda företag är därför minst lika viktig som miljöbetygen. Betygen för vissa företag kan bli fel av uppenbara anledningar men denna typ av fel är lätta att identifiera. Därför är det viktigt att tolka enskilda resultat med försiktighet och mot bakgrund av hur metoden fungerar. Det går inte att ersätta ett komplext och mångfacetterat fenomen med ett enda måttal utan att förlora information, inte ens med maskininlärningsalgoritmer som beaktar alla 100+ indikatorernas värden för miljöbetyget.

Slutligen finns ett litet utdrag ur de tre komponenterna av HBR nedan i Tabell 7. Tabellen är sorterad efter det sociala betyget, varför alla företag i tabellutdraget har nästan samma sociala betyg. Noterbart är att ett företag kan ha olika riskexponering i de tre komponenterna.

Tabell 7: Utdrag ur de tre komponenterna av HBR.

Company Data					Environment				Social				Governance			
company name	Country	Press freedom	Market Cap [\$B]	GICS Sector Name	Proposed env rating	Conservative env rating	E-score	# env controversies	Proposed soc rating	Conservative soc	S-score	# soc controversies	Proposed gov rating	Conservative gov rating	G-score	# gov controversies
Telia Company AB	SE	91	17.4	Communication Serv	66	65	72	0	9	8	81	13	51	50	68	0
Freeport-McMoRan Inc	US	76	13.7	Materials	1	0	75	1	10	5	77	15	28	23	76	1
Arconic Inc	US	76	10.8	Industrials	2	0	92	2	10	5	86	13	48	43	54	1
MAN SE	DE	88	7.2	Industrials	14	13	81	0	10	9	81	9	94	93	39	1
Eli Lilly and Co	US	76	99.4	Health Care	16	9	75	0	10	3	88	27	6	0	71	0
UnitedHealth Group Inc	US	76	199.1	Health Care	31	24	73	0	10	3	73	20	13	6	80	1
Visa Inc	US	76	345.3	Information Technol	31	23	53	0	10	2	60	26	4	0	40	0

Referenser

- Berg, F., Kölbl, J., Rigobon, R., 2019. Aggregate Confusion: The Divergence of ESG Ratings. SSRN Electronic Journal.
- Chatterji, A., Levine, D., 2006. Breaking down the wall of codes: Evaluating non-financial performance measurement. *California Management Review* 48, 29–51.
- Chatterji, A.K., Durand, R., Levine, D.I., Touboul, S., 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal* 37, 1597–1614.
- Chen, C.M., Delmas, M., 2011. Measuring corporate social performance: An efficiency perspective. *Production and Operations Management* 20, 789–804.
- Delmas, M., Blass, V.D., 2010. Measuring Corporate Environmental Performance: The Trade-Offs of Sustainability Ratings. *Business Strategy and the Environment* 19, 245–260.
- Delmas, M.A., Etzion, D., Nairn-Birch, N., 2013. Triangulating environmental performance: What do corporate social responsibility ratings really capture? *Academy of Management Perspectives* 27, 255–267.
- Hillman, A.J., Keim, G.D., 2001. Shareholder value, stakeholder management, and social issues: What's the bottom line? *Strategic Management Journal* 22, 125–139.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2019. Explainable AI for Trees: From Local Explanations to Global Understanding. arXiv.
- Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions, in: *Advances in Neural Information Processing Systems*. Neural information processing systems foundation, pp. 4766–4775.
- Nofsinger, J.R., Sulaeman, J., Varma, A., 2019. Institutional investors and corporate social responsibility. *Journal of Corporate Finance* 58, 700–725.
- Orlitzky, M., Schmidt, F.L., Rynes, S.L., 2003. Corporate social and financial performance: A meta-analysis. *Organization Studies* 24, 403–441.
- Semenova, N., Hassel, L.G., 2015. On the Validity of Environmental Performance Metrics. *Journal of Business Ethics* 132, 249–258.
- Shmueli, G., 2010. To explain or to predict? *Statistical Science* 25, 289–310.
- Trumpp, C., Endrikat, J., Zopf, C., Guenther, E., 2013. Definition, Conceptualization, and Measurement of Corporate Environmental Performance: A Critical Examination of a Multidimensional Construct. *Journal of Business Ethics* 126, 185–204.
- Wong, C., Petroy, E., 2020. Rate the Raters 2020: Investor Survey and Interview Results.