

EN NY GENERATION FÖREBYGGANDE AV ARBETSOLYCKSFALL MED MASKININLÄRNING

Sammanfattning

**May Shayboun, Christian Koch, & Dimosthenis
Kifokeris**

2022-04-19

SAMMANFATTNING

Bakgrund

Säkerheten på byggarbetsplatser är fortfarande ett område som väcker stort intresse inom byggsektorn. Antalet arbetsplatsolyckor har minskat i Sverige men bara till en nivå som är konstant under de senaste tio åren. Även om Sverige visar sig prestera bättre jämfört med andra europeiska länder, så är byggbranschen fortfarande ansvarig för en femtedel av antalet arbetsrelaterade dödsolyckorna i Europa. Behovet av att minska antalet dödsfall till följd av arbetsolyckor i byggsektorn har aldrig varit större eller mer påtaglig. I Sverige har flera initiativ genomförts för att minska frekvensen av olycksfall, dock så harriskanalysmodeller och orsakssamband visat sig vara sällsynta i detta sammanhang.

Rapporteringen av arbetsplatsrelaterade olyckor och tillbud genererar stora mängder data som potentiellt kan bidra till att få fram de bakomliggande orsakerna. Vidare ses ett ökat intresse inom forskningsvärlden att med hjälp av maskininlärning analysera olycksfallsdata. State-of-the-art forskningsinsatser arbetar med att tillämpa maskininlärning för att analysera textdata från ackumulerade olycksrapporter, identifiera bidragande faktorer och ta fram olycksinformation. Sådana lösningar som är skapade av maskininlärningsmodeller skulle kunna leda till förändringar både för enskilda företag och branschen som helhet. Som en del av processen för maskininlärning modellering så inkluderade vi feedback från branschen och domänexperterna för att anpassa prototypen mot industrins efterfrågan.

Syfte

Syftet med denna avhandling är att undersöka hur metoder och tekniker för maskininlärning skulle kunna tillämpas för att utveckla en forskningsbaserad prototyp för att förebygga arbetsolyckor i ett entreprenadföretag. Uppsatsens fokus ligger på att utforska en utvecklingsprocess som överbryggar maskininlärning och dataanalys med säkerhetskontexten i ett entreprenadföretag. Avhandlingen bygger på olycksfallsmodeller (Accident Causation Models, ACM) och maskininlärningsmetoder genom tillämpningen av Cross Industry Standard Process Development Method (CRISP-DM). De sex faserna i CRISP-DM inkluderar affärsförståelse (business understanding), dataförståelse (data understanding), dataförberedelse (data preparation), modellering (modelling), utvärdering (evaluation), och utveckling (deployment), och visas i Figur 1 (se nästa sida).

Olycksfallsmodeller, maskininlärningsmetoder och CRISP-DM användes för att tolka och förstå det empiriska materialet i olycksrapporter och som komplement så intervjuades berörda personer inom enheten hälsa och säkerhet (H&S).

Genomförande

Detta licentiatarbete bidrar genom sina resultat till att överbrygga gapet mellan maskininlärning och dataanalys genom tillämpad kunskap om säkerhetskontexten i ett entreprenadföretag. Arbetet har utvecklat en metod för selektering av lämplig algoritm utifrån erforderliga kriterierna för maskininlärningsmodellen. Det beslutades att tillämpa CRISP-DM som metod för att förstå sammanhanget och samla in kunskap om potentiell användning av maskininlärning utifrån ett affärsperspektiv. ACM var också väsentliga i tolkning med maskininlärningsmodellen, särskilt för att identifiera komponenter i olycksanalys och kategorier av orsaker.

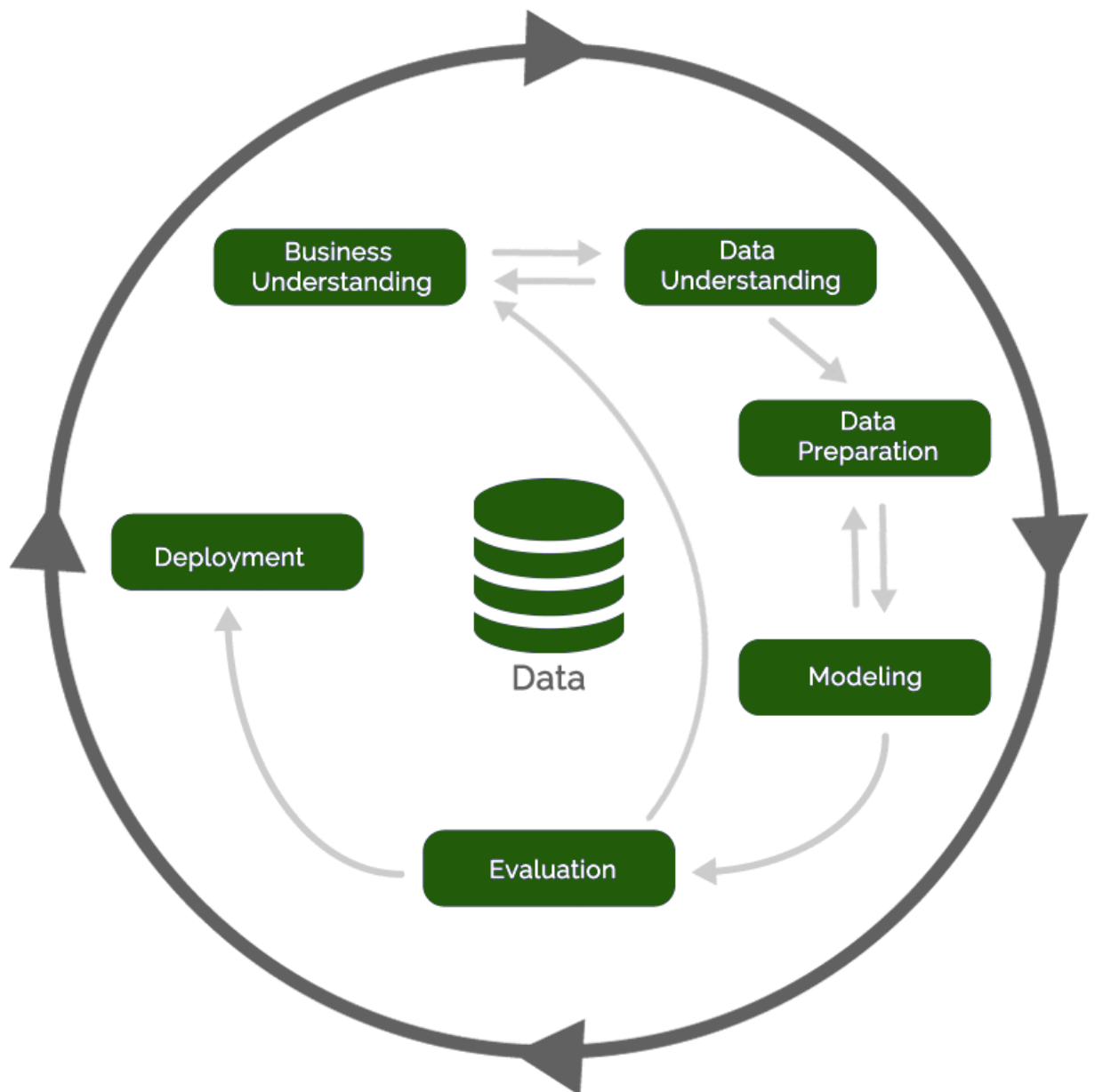


Fig. 1. Sex faserna av CRISP-DM

Arbetet med olycksorsaksanalys har resulterat i flera klassificeringar, både linjära och olinjära modeller enligt det antagna logiska händelseförloppet som leder från risk

till olyckor (t.ex., Bow-tie modellen). Andra klassificeringar finns också, men ACM har delats in i grupper utifrån olika stadier och orsakssamband. De enkla linjära modellerna tillskrev olyckor fysiska/mekaniska och mänskliga fel. ACM blev sedan förknippad med komplexa linjära modeller eftersom de alltmer övervägde den interaktion som sker mellan latent organisatoriska faktorer och osäkra beteenden. Komplexa icke-linjära modeller uppmuntrade en bredare syn på systemrelaterade faktorer som svar på den växande komplexiteten och stramare kopplingar inom industriella domäner. De nyare och mer avancerade orsaksmodeller förklarar olyckor som orsakades av den dynamiska och icke-linjära sambandet mellan flera faktorer inom hela systemet, inklusive politiska och regulativa faktorer.

ACM utvecklades till att inkludera högre nivåer av orsakssamband. Dessutom anerkändes stokasticitet i olyckans svårighetsgrad. Beteende och avancerade socio-tekniska och kulturella modeller användes i relevant domänlitteratur i byggforskningssammanhang, medan de systembaserade modellerna väldigt sällanhar tillämpats inom byggbranschen. Det fortsätter att inträffa olyckor i byggbranschen och det finns ett behov av att undersöka teorier och modeller för olycksorsaksbildning som utveckling av den kvantitativa dataanalys som finns nu härrör från många registrerade olyckor.

Den maskininlärningsbaserade metoden för olycksanalys inkluderar övervakad, oövervakad och semiövervakad maskininläring. Oövervakat lärande är en metod för datautforskning eller beskrivning som används där inga specifika förtilldelade etiketter (label) för ingångs- eller utgångsfunktionerna har tilldelats. I jämförelse, bygger övervakad maskininläring på att en redan märkt ingång (etiket) kartläggs till utdata. Semi-övervakad maskininläring består av en kombination av de två andra tillvägagångssätten. Övervakade maskininläring-algoritmer kan ytterligare kategoriseras i linjära och icke-linjära algoritmer, var och en med olika egenskaper, styrkor och svagheter. Algoritmerna organiserades utifrån deras egenskaper (t.ex., tolkningsbarhet, noggrannhet, generaliserbarhet).

Denna avhandling har använt ett övergripande kvalitativt-tolkande reflexiv metodologiskt tillvägagångssätt. Detta tillvägagångssätt kombinerade olika tolkningsnivåer. ACM och maskininläring olycksanalys valdes som de huvudsakliga teoretiska ramarna för att besvara en övergripande forskningsfråga och fyra relaterade delfrågor.

Den tillhörande empiriska forskningen omfattade insamling av olycksrapporter och intervjuer bedrivs inom H&S-enheten i ett entreprenadföretag. CRISP-DM och maskininläring-algoritmerna var anställd för att utveckla och analysera en tillämpad maskininlärningsmodell. ACM, maskininlärningsalgoritmer och CRISP-DM gav den önskade mångfalden i tolkningen av det empiriska materialet. Den

praktiska forskningsmetoden bestod av fyra sekventiella steg, inklusive tre artiklar och den maskininlärning-baserade analysen av olycksdata.

Maskininlärningsbaserad olycksanalys kan skapa kunskap om olyckor i ett entreprenadföretag eller annat organisationer, såsom föremål och kombinationer av situationer som orsakar olyckor. Litteraturen i detta sammanhang visade olika sätt genom vilka tillämpningen av maskininlärning-algoritmer kan förbättra kunskapen om olyckor. maskininlärning-modeller kan användas vid uppskattning av svårighetsgrad, klassificering av olyckstyper, informationsutvinning och generering av scenarier för säkerhetsutbildning. Men litteraturen visar också att utvinning av ny kunskap om olyckor i ett entreprenadföretag hindrades av en mängd utmaningar. Främst systematiseringen av maskininlärning-processen, genomförbarheten av implementering och fokusera på förutsägelse av svårighetsgrad.

Den granskade litteraturen om tillämpad maskininlärning i analys av olycksrapporter indikerade behovet av standardisering av utvecklingsprocessen vad gäller genomförbarheten av implementering och utvärdering.

Men implementeringen av CRISP-DM som en process tillförde viktiga komponenter till förståelsen av kontexten, såsom antaganden om säkerhetsprocesser och förebyggande av olyckor. Även om CRISP-DM befanns vara för allmänt för att ge specifika riktlinjer för maskininlärning-prototyputveckling, det gav en ryggrad för applikationsdomänen.

Resultat och slutsatser

Resultatet av avhandlingen visade att man genom en analys av olycksrapporter med hjälp av maskininlärning kan uppnå ny kunskap om olyckor. Det fanns dock flera utmaningar som visade sig hindra utvinning av kunskap och tillämpning av maskininlärning. De identifierade utmaningarna relaterade främst till standardisering av utvecklingsprocessen och genomförbarheten av implementering och utvärdering. Dessutom är inte tendensen hos den maskininlärning-relaterade litteraturen att fokusera på att förutsäga svårighetsgrad av olycksfall förenlig varken med maskininlärning-analysens funktion eller resultat från olycksorsakslitteratur som betraktar svårighetsgrad som ett stokastiskt element. Analysen drog vidare slutsatsen att ACM verkade ha nått ett moget stadium, där ett nytt tillvägagångssätt behövs för att förstå de regler som styr relationer mellan framväxande nya risker – snarare än systematiseringen av själva riskerna. Analys av olycksrapporter av maskininlärning behöver ytterligare forskning i systematiserade metoder för sådan analys i byggbranschens domän och i entreprenadföretagens sammanhang – som endast få forskningsinsatser har fokuserat när det gäller maskininlärning-utvärderingsmått och förbearbetning av data.

Författarens erkännanden

Författarna vill tacka SBUF för finansiering och kontinuerligt stöd till detta projekt.
Detta forskningsarbete hade inte kunnat slutföras annars.