

Henrik Grönblom

TYÖTAPATURMIEN ENNUSTAMINEN RAKENNUSTEOLLISUUDESSA

Diplomityö
Johtamisen ja talouden tiedekunta
Sari Tappura
Jari Turunen
Toukokuu 2023

TIIVISTELMÄ

Henrik Grönblom: Työtapaturmien ennustaminen rakennusteollisuudessa
Diplomityö
Tampereen yliopisto
Johtamisen ja talouden tiedekunta
Toukokuu 2023

Työturvallisuuteen keskitytään yhä enemmän ja samalla suojavälineiden käyttö on lisääntynyt sekä suojainten laatu on parantunut. Rakennusteollisuus on yksi riskialtteimmista aloista Suomessa, vaikkakin työpaikkatapaturmien määrät ovat olleen alalla laskusuunnassa. Laskentatavasta riippuen ala on jopa suurin työllistäjä Suomessa, joten työturvallisuuden kehittämiseen ja seuraamiseen on hyvä panostaa.

Tämä tutkimus liittyy organisaatiossa käytettävien ennustemallien kehittämiseen. On huomattu, että nykyisin käytössä olevaa ennustemallia paremmin voitaisiin ennustaa mahdollisesti toisella ennustemenetelmällä ja käyttäen erilaista selittävää tekijää. Tämän diplomityön tarkoituksena oli tutkia neuroverkkojen tuottaman ennustemallin hyödyllisyyttä työtapaturmien vuosittaisen ennustamisen tukena.

Tutkielma koostuu kirjallisuuskatsauksesta työtapaturmiin ja niihin liittyvään lainsäädäntöön sekä uusien esiteltävien ennustemallien hyödyntämiseen. Aiemmin saatujen havaintojen pohjalta valittiin tiettyjä selittäviä tekijöitä mukaan ennustemalliin, joilla voisi olla mielekästä ennustaa työtapaturmia. Tutkimusaihe on rajattu rakennusalan työpaikkatapaturmiin lukuun ottamatta maa- ja vesirakentamisen toimialaa.

Tutkimuksessa selvitetään, miten LSTM- ja GRU-neuroverkot soveltuvat työpaikkatapaturmien ennustamiseen. Samassa yhteydessä laaditaan 12 kuukauden ennuste LSTM- ja GRU-neuroverkkojen avulla sattumisvuodelle 2022. Tilastovuosi 2022 valmistuu vasta myöhemmin kesällä 2023, joten tulosten tarkempi validointi ei vielä tutkimuksen tekoaikaan ole mahdollista.

Tutkimuksen tuloksena LSTM-malli ennustaa 13095 ja GRU-malli 13304 työtapaturmaa rakennusteollisuuden alalle vuonna 2022. Viimeisen kymmenen vuoden aikana on sattunut keskimäärin 12 859 vahinkoa ja viimeisen viiden vuoden aikana on sattunut keskimäärin 12973 vahinkoa. Laadittujen ennusteiden voidaan todeta olevan linjassa historiassa toteutuneiden vahinkomäärien kanssa. LSTM-mallin tuottama ennuste on hieman lähempänä historiallisia vahinkojen lukumäärien keskiarvoja ja vaikuttaa näin ollen tuottavan käytetyistä malleista paremman ennusteen. Historiaan peilaten käytettävien ennustemallien voidaan todeta olevan sopivia käytännön työkaluja vuosittaisen ennusteen laskemiseksi. Työssä laadittuja ennustemenetelmiä hyödyntäen tullaan laatimaan työtapaturmien ennuste syksyllä 2023.

Avainsanat: Ennustaminen, työtapaturmat, neuroverkot, LSTM, GRU.

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ABSTRACT

Henrik Grönblom: Predicting accidents at work in the construction industry
Master's Thesis
Tampere University
Degree Programme in Management and Information Technology
May 2023

Work safety is increasingly being focused on, and at the same time the use of protective equipment has increased and the quality of protective equipment has improved. Construction industry is one of the most high-risk industries in Finland. Depending on calculation the industry is the biggest employer in Finland, so it is good to invest in the development and monitoring of work safety.

The purpose of this study is related to the development of prediction models used in the organization. It has been noticed in the organization that the currently used forecasting model could possibly be better forecasted with another forecasting method and by using a different explanatory factor. The aim of this study was to investigate the usefulness of the prediction model produced by neural networks in support of the annual prediction of occupational accidents.

The thesis consists of a literature review on occupational accidents and related legislation as well as the utilization of the new prediction model. Based on earlier observations, certain explanatory factors were selected to be included in the prediction model which could be meaningful in predicting occupational accidents. The research topic is limited to workplace accidents in the construction industry, without civil engineering industry.

The research examines how LSTM and GRU neural networks are suitable for predicting occupational accidents. In the same context, a 12-month forecast is prepared using LSTM and GRU neural networks for the event year 2022. The statistical year 2022 will not be completed until later in the summer of 2023, so a more precise validation of the results is not yet possible at the time of the study.

As a result of the research, the LSTM model predicts 13095 and the GRU model 13304 occupational accidents in the construction industry in 2022. During the last ten years, an average of 12859 injuries have occurred, and during the last five years, an average of 12973 injuries have occurred. The prepared forecasts can be said to be in line with the accidents that have occurred in history. The forecast by the LSTM model is slightly closer to the historical averages of the accidents and thus seems to produce a better forecast than the GRU model. Based on historical data, the forecast models used seem to be suitable practical tools for calculating the annual forecast. The created forecast model of occupational accidents will be used in the fall of 2023.

Keywords: Forecasting, Occupational Accidents, Neural Networks, LSTM, GRU.

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

ALKUSANAT

Tämä diplomityö on tehty opinnäytetyönä Tampereen yliopiston Johtamisen ja tietotekniikan diplomi-insinöörin tutkinto-ohjelmaan sekä toimeksiantona Tapaturmavakuutuskeskukselle. Työn tarkistajina ovat toimineet yliopisto-opettaja Sari Tappura sekä yliopistonlehtori Jari Turunen. Päädyin tähän aiheeseen ohjaajani Jari Turusen opastuksella.

Tahtoisin kiittää ohjaajia Sari Tappuraa sekä Jari Turusta työn valmiiksi saattamisessa. Olen saanut tukea ja työn perään on kyselty sopivissa määrin, jotta valmistuminen tässä kaiken kiireen keskellä on mahdollista. Kiitän myös Tapaturmavakuutuskeskuksen henkilöstöä vinkeistä ja aineistosta tutkimuskäyttöön.

Erityiskiitokset haluan välittää työtoverilleni Janne Jylhälle, jonka vankkumaton tuki ja kannustus auttoivat korvaamattomasti itselleni uuden ja vaikean aihealueen käsittelemisessä. Lisäksi kiitokset kotiväelle, että ovat jaksaneet kannustaa ja toimia tukena.

Porissa, 17.5.2023

Henrik Grönblom

SISÄLLYSLUETTELO

1. JOHDANTO	1
1.1 Tutkimuksen tausta	2
1.2 Tutkimuksen tavoitteet, rajaus ja tutkimusongelmat	2
2. TEOREETTINEN TAUSTA	4
2.1 Työtapaturma- ja ammattitautivakuutus.....	4
2.2 Tapaturmavakuutuskeskus ja sen tehtävät.....	5
2.3 Työtapaturmat Suomessa	7
2.3.1 Työtapaturmat rakennusteollisuudessa	8
3. ENNUSTEMENETELMÄT	12
3.1 Vektoriautoregressiivinen malli (VAR-malli).....	12
3.2 Koneoppiminen	13
3.3 Neuroverkot	15
3.3.1 LSTM-arkkitehtuuri.....	18
3.3.2 GRU-arkkitehtuuri	19
4. TYÖTAPATURMIEN MÄÄRÄN ENNUSTAMINEN.....	21
4.1 Ongelman asettelu	21
4.2 Aineisto	21
4.2.1 Työpaikkatapaturmat	23
4.2.2 Aloitetut rakennushankkeet.....	24
4.3 Mallien soveltaminen aineistoon.....	25
5. TULOKSET	30
5.1 LSTM-malli.....	30
5.2 GRU-malli	33
5.3 Mallien vertailu	36
6. JOHTOPÄÄTÖKSET	37
7. LÄHTEET	39
LIITE A: TYÖPAIKKATAPATURMAT SATTUMISVUOSITTAIN	41
LIITE B: LISTAUS AKTIVAATIOFUNKTIOISTA	42
LIITE C: LSTM-NEUROVERKKO.....	43
LIITE D: GRU-NEUROVERKKO	46

LYHENTEET JA MERKINNÄT

AI	Artificial Intelligence
CEC	Constant Error Carousel
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
RPA	Robotic Process Automation
RMSE	Root Mean Square Error
TVK	Tapaturmavakuutuskeskus
VAR	Vector Autoregression

1. JOHDANTO

Työtaturma- ja ammattitautivakuutus turvaa työntekijälle korvaukset työtaturman ja ammattitaudin varalta. Korvaukset voivat olla suuria. Suurin osa korvauksista on työkyvyttömyydestä aiheutuvia ansionmenetykskorvauksia tai sairaanhoidosta aiheutuneita kustannuksia. Lakisääteinen työtaturma- ja ammattitautivakuutus perustuu oikeuteen saada korvausta työtaturmasta ja ammattitaudista riippumatta siitä onko työnantaja huolehtinut vakuuttamisesta vai ei. Oikeus korvauksen saamiseen perustuu työ- tai virkasuhteen olemassaoloon.

Tapaturmavakuutuskeskus (TVK) tuottaa luotettavia tilastoja työtaturma- ja ammattitautivahinkojen kehittymisestä. Samaan aikaan halutaan tuottaa ennusteita tulevasta. Rakentaminen ja rakennetun ympäristön ala työllistää noin viidenneksen Suomen työllisistä. Rakennusteollisuus työllistää n. 250 000 henkilöä ja kiinteistöhoito sekä alaan liittyvät muut palvelut huomioiden yli puoli miljoonaa. Näin laskettuna ala on suurin työllistäjä Suomessa. Rakennusteollisuuden turvallisuudessa on tehty merkittäviä parannuksia viimeisen kymmenen vuoden aikana. Nykyisellään perusasiat, kuten henkilökohtaiset suojaimet ovat kunnossa, mutta vielä on kehittämistä työturvallisuudessa. Joidenkin vertailujen mukaan rakennusalan työturvallisuus on heikompi Suomessa kuin esimerkiksi naapurimaa Ruotsissa. Tilastojen vertailu on hankalaa tilastoinnin erojen vuoksi, mutta kuolemantapausten suhteuttaminen tehtyihin työtunteihin kääntää työturvallisuustilanteen kuitenkin päinvastaiseksi. Kuolemantapausten tilastoinnissa ei voi olla tulkinnanvaraisuutta eri maiden välillä. (Rakennusteollisuus ry, 2023a; Rakennusteollisuus ry, 2023b)

Rakennusteollisuus on hyvin suhdanneherkkä ala ja maailmantalouden tilanne vaikuttaa väistämättä rakennusteollisuuden näkymiin. Energian hinnan nousu kasvattaa taantumariskiä ja samalla keskuspankit tekevät rahapoliittisia toimia vähentääkseen kokonaiskysyntää ja inflaatiota. Rakentamisen kustannukset ovat nousseet suuresti ja materiaalien saatavuudessa on ollut viivästymisiä. Kaikkea maailmantilanteen aiheuttamaa viivästystä ei osata ottaa huomioon aikataulutuksessa, jolloin aikatauluista jäädään jälkeä, syntyy kiirettä ja tapaturmariski kasvaa (Työterveyslaitos, 2019).

Työ tehdään Tapaturmavakuutuskeskuksen toimeksiannosta. Tapaturmavakuutuskeskus koordinoi ja kehittää työtapaturma- ja ammattitautivakuutuksen toimeenpanoa Suomessa ja käsittelee korvausasiat vahinkotapahtumissa, jotka ovat sattuneet vakuutamattomassa työssä. Tämän lisäksi Tapaturmavakuutuskeskus tilastoi Suomessa tapahtuneet palkansaajien työpaikkatapaturmat ja ylläpitää tilastorekistereitä työtapaturma- ja ammattitautivakuutuksen vakuutus-, vahinko-, pääoma- ja korvaustiedoista.

1.1 Tutkimuksen tausta

Rakennusteollisuuden työturvallisuus on kehittynyt parempaan suuntaan viimeisen vuosikymmenen aikana. Etenkin vakavien tapaturmien määrä on vähentynyt, vaikka tehdyn työn määrä tunneissa mitattuna on kasvanut. Lievät tapaturmat ovat pysyneet suunnilleen samoissa lukumäärissä. Lievien tapaturmien esiintyvyys yleensä kasvaa rakentamisen noususuhdanteen aikana, mutta vakavien tapaturmien ei. (Työterveyslaitos, 2019)

Rakennusalan työturvallisuuskehitys koostuu kolmesta päätekijästä:

1. työturvallisuuskulttuurin muutos
2. suojavälineet
3. panostaminen työturvallisuuteen.

Nykyisin työturvallisuus otetaan vakavammin kuin ennen. Samalla suojavälineiden käyttö on lisääntynyt huomattavasti ja niiden laatu on parantunut. Lisäksi rakennusalan yritykset panostavat huomattavasti työturvallisuuteen. Työturvallisuus ymmärretään nykyään niin taloudellisesta kuin yrityksen maineen kautta. Rakennusalan työturvallisuuden kehitystä voidaan mitata erilaisin tavoin. Mahdollisimman luotettava ennustaminen työtapaturmien kehityksestä auttaa laajalti muun muassa riskienhallinnan toimenpiteitä suunniteltaessa ja myös talouden tilastojen tulkinnassa. (Työterveyslaitos, 2019)

1.2 Tutkimuksen tavoitteet, rajaus ja tutkimusongelmat

Tutkimusaihe rajataan rakennusalan työpaikkatapaturmiin. Työssä kuvataan lyhyesti, mitä koneoppiminen on. Tämän jälkeen esitellään aikasarjan ennustamiseen sopivia koneoppimismenetelmiä ja tutustutaan yleisellä tasolla organisaatiossa käytettyyn VAR-ennustemalliin. Tutkimuksen tarkoituksena on tuottaa ennuste 12 seuraavan kuukauden aikana sattuvien tapaturmien lukumäärästä LSTM- ja GRU-mallien avulla. Tapaturmavakuutuskeskuksessa on käytössä vahinkojen ennustamiseen aiemmin toteu-

tettu VAR-malli. VAR-malli on lineaarinen, kun taas neuroverkkojen avulla voidaan hyödyntää epälineaarisen mallin mukanaan tuomia hyötyjä.

Työ keskittyy kahteen ennalta määriteltyyn tutkimusongelmaan ennustemallien käytöstä:

1. Miten LSTM- ja GRU-neuroverkot soveltuvat työpaikkatapaturmien lyhyen aikavälin ennustamiseen?
2. Millaiset tapaturmaennusteet neuroverkoilla saadaan sattumisvuodelle 2022?

Työ keskittyy ainoastaan neuroverkkojen soveltuvuuteen ja käyttöön työpaikkatapaturmien lyhyen aikavälin ennustamisessa Tapaturmavakuutuskeskuksen työturvallisuusyksikössä. Tapaturmavakuutuskeskus julkaisee syksyisin arvion koko edeltävän vuoden aikana sattuneista työpaikkatapaturmista.

Tässä tutkielmassa on tarkoitus käsitellä syväoppivia neuroverkkoja ja niiden sovelluksia. Työssä esitetään neuroverkot, sekä tarkemmin LSTM- sekä GRU-neuroverkkorakenteet. Lisäksi käydään tarvittavilta osin koneoppimisen perusteita. Lopuksi sovelletaan edellä mainittuja neuroverkkorakenteita työtapaturmavahinkojen lyhyen aikavälin ennustamiseksi.

2. TEOREETTINEN TAUSTA

Tarve suojautua sairauden, vanhuuden tai vahinkojen varalta on aina ollut olemassa. Keskinäisen vastuunkannon tilalle syntyi tahoja kantamaan toisten puolesta riskiä korvausta vastaan. Nykyisen muotoinen vakuutusliiketoiminta pohjautuu tähän samaan toimintatapaan. (Rantala, 2009) Tapaturma- ja ammattitautilaki (TyTAL) on osa sosiaalivakuutusta ja oikeuttaa työntekijän saamaan korvausta työtapaturmasta tai ammattitaudista työ- tai virkasuhteessa. (TyTAL, 2015)

Tavoitteena on avata työtapaturma- ja ammattitautivakuutusta yleisesti sekä siihen liittyvää lainsäädäntöä lyhyesti. Samassa yhteydessä esitetään Tapaturmavakuutuskeskus ja sen tehtävät sekä taustoitetaan tutkimuksen taustalla olevaa aineistoa.

2.1 Työtapaturma- ja ammattitautivakuutus

Työtapaturma- ja ammattitautivakuutus korvaa Suomessa noin 130 000 työtapaturmaa ja ammattitautia vuosittain. Työnantajan tulee ottaa vakuutus yksityisestä vakuutusyhtiöstä, mutta yrittäjälle vakuutus on vapaaehtoinen sillä ehdolla, että yrittäjällä on yrittäjän eläkelain mukainen vakuutus. Mikäli kyseessä on yksityisen tai kunnallisen sektorin työnantaja, niin työnantajan tulee ottaa työtapaturma- ja ammattitautivakuutus yksityiseltä vakuutusyhtiöltä. Valtiolla ei ole vakuuttamisvelvollisuutta, mutta korvaukset hoidetaan keskitetysti Valtiokonttorin toimesta. (Salo, 2015)

Työtapaturma käsitteenä koostuu kolmesta osasta, mitkä yhdessä muodostavat työtapaturman määritelmän: tapaturma, vahinko syy-yhteydessä tapaturmaan sekä tapaturman sattumisolosuhteet. Työtapaturma määritellään työntekijän kohtaamana tapaturmana, jos se aiheuttaa vamman tai sairauden. Lisäehtoina on tapaturman sijaintiin liittyviä asioita useampia. Tapaturma on sattunut työssä, työntekopaikan alueella tai sen ulkopuolella työmatkalla tai jossakin erityisissä olosuhteissa, kuten koulutustilaisuudessa tai muussa erityisessä tilaisuudessa. (Salo, 2015)

Työtapaturma- ja ammattitautivakuutuksen tarkoituksena on taata työntekijälle korvaukset työtapaturmasta ja ammattitaudista. Työtapaturma- ja ammattitautivakuutus on osa suomalaista sosiaaliturvaa ja sen vuoksi sitä kutsutaan myös sosiaalivakuutukseksi.

si. Sosiaalivakuutus on etukäteen maksettavista vakuutusmaksuista johtuen erilainen muuhun sosiaaliturvaan verrattuna. Myös työeläke, sairausvakuutus, kansaneläke ja työttömyysturva ovat sosiaalivakuutuksia. Työtapaturma- ja ammattitautivakuutus on sidoksissa työsuhteeseen työeläkkeen tavoin ja poikkeaa täten muusta sosiaalivakuuttamisesta. Työtapaturma- ja ammattitautivakuutus on Suomen vanhin sosiaalivakuutus. Liikennevakuutus ja potilasvakuutus ovat myös lakisääteisiä vakuutuksia. (Salo, 2015)

Työtapaturma- ja ammattitautivakuutuksen tarkoituksena on taata sosiaaliturva työtapaturman ja ammattitaudin jälkeen. Työtapaturma- ja ammattitautivakuutuksessa on myös oma erillinen muutoksenhakujärjestelmänsä. Ensimmäisenä asteena toimii tapaturma-asioiden muutoksenhakulautakunta ja toisena vakuutusosoikeus. Korvauspäätöksestä voi hakea myös valituslupaa korkeimmalta oikeudelta. (Salo, 2015)

Työtapaturma- ja ammattitautivakuutus on ensisijainen vakuutus. Ensisijaisuudella tarkoitetaan sitä, että jos kyseessä on työtapaturma- ja ammattitautilain mukaisesti korvattava työtapaturma tai ammattitauti niin korvaus maksetaan ensisijaisesti tästä vakuutuksesta. Ensisijaisuus tarkoittaa samalla myös sitä, että lisäkorvausta on mahdollista saada muista vakuutuksista, jos ensisijainen vakuutus, eli tässä tapauksessa työtapaturma- ja ammattitautivakuutus, on korvauksiltaan pienimpi. Parhaimpana esimerkkinä tästä toimii esimerkiksi työpaikan ja asunnon välisellä työmatkalla tapahtunut liikennevahinko. Ensisijaisesti vahinko korvataan työtapaturma- ja ammattitautivakuutuksesta, mutta tilannekohtaisesti lisäkorvausta voidaan hakea liikennevakuutuksesta. Tapaturmavakuutuskeskus hoitaa pientyönantajien työssä sattuneet vahinkotapahtumat sekä laiminlyöntitilanteet. Maatalousyrittäjien eläkelaitos Mela hoitaa maatalousyrittäjien sekä tieteen ja taiteen apurahansaajien vakuutukset. Työtapaturma- ja ammattitautivakuutuksen toimeenpano vakuutusyhtiössä rinnastetaan viranomaisen toimintaan, jolloin siinä sovelletaan lisäksi myös muun muassa hallintolain säännöksiä. Tiivistetysti voidaan siis sanoa, että yksityiset vakuutusyhtiöt toteuttavat työtapaturma- ja ammattitautivakuutuksen toimeenpanossa julkista hallintotehtävää. (Salo, 2015)

2.2 Tapaturmavakuutuskeskus ja sen tehtävät

Tapaturmavakuutuskeskus ylläpitää rekisteriä työnantajista, joilla on lakisääteinen vakuutus työntekijöitään varten. Työtapaturma- ja ammattitautivakuutusta tarjoavan yhtiön tulee olla Tapaturmavakuutuskeskuksen jäsen. (TyTAL, 2015)

Tapaturmavakuutuskeskuksen tehtävä on toimia työtapaturma- ja ammattitautivakuutuksen toimeenpanemisessa ja kehittämisessä. Tapaturmavakuutuskeskuksen yleisinä tehtävinä on

- lain yhdenmukaisen toimeenpanon edistäminen sekä toimeenpanojärjestelmän kehittäminen
- lain toimeenpanoa koskevien asioiden tiedottaminen sekä yhteistyön edistäminen
- vakuutuslajin vakuutusyhtiöiden sekä viranomaisten yhteistyön edistäminen
- yhdenmukaistaa lain toimeenpanoa antamalla suosituksia. (TyTAL, 2015)

Tapaturmavakuutuskeskuksen tehtävänä on hoitaa niin sanottua korvaustakuuta. Korvaustakuu tulee kysymykseen silloin, jos vakuutusyhtiö laiminlöisi korvausasian selvittämisen, päätöksen antamisen tai korvauksen maksamisen. Viivästymistilanteissa korvausta voidaan hakea Tapaturmavakuutuskeskukselta viivästyneiltä osin. Tapaturmavakuutuskeskus perii aiheutuneet kustannukset jälkikäteen tehtävän laiminlyöneeltä vakuutusyhtiöltä. (Salo, 2015)

Työtapaturma- ja ammattitautivakuutuksen vakuuttamisen yleisvalvonta kuuluu Tapaturmavakuutuskeskukselle. Valvontaa suoritetaan yhteistyössä muun muassa Eläketurvakeskuksen sekä Työllisyysrahaston kanssa, joilla molemmilla on vastaava valvontatehtävä. Lisäksi yhteistyötä tehdään työsuojeluviranomaisen kanssa. Valvonnan suorittamisessa käytetään vakuutusrekisteriä. Vakuutusrekisteri sisältää tiedon niistä työnantajista, jotka ovat ottaneet pakollisen vakuutuksen työntekijöilleen. Tiedot vakuutusrekisteriin saadaan vakuutusyhtiöiltä. Mikäli työnantaja on laiminlyönyt vakuutusvelvollisuutensa, niin se on velvollinen maksamaan vakuutusmaksua vastaavan maksun sekä laiminlyöntimaksun. Vakuutusmaksua vastaava maksu voi koostua kulumassa olevalta ja viideltä viimeiseltä kalenterivuodelta. Laiminlyöntimaksun suuruus on maksimissaan kolme kertaa vakuutusmaksua vastaavan maksun suuruinen. Mikäli havaittua laiminlyöntiä ei korjata annetussa määräajassa niin Tapaturmavakuutuskeskus ottaa kyseiselle työnantajalle niin sanotun pakkovakuutuksen. Laiminlyöntitapauksissa Tapaturmavakuutuskeskus tekee hakemuksen Valtiokonttorille edellä mainittujen maksujen määräämiseksi. Vahingon sattuessa työntekijän korvausasia käsitellään Tapaturmavakuutuksessa samalla tavalla, kuin vahinko olisi sattunut vakuutetussa työssä. Tällaisessa tapauksessa vakuuttamisen laiminlyönyt työnantaja joutuu osallistumaan etuuk-sien maksamiseen olemassa olevien rajoitteiden puitteissa. (Salo, 2015)

Työpaikkaonnettomuuksien tutkinnasta käytetään nimitystä TOT-tutkinta. TOT-tutkinnan johtamisesta ja valvonnasta vastaa Tapaturmavakuutuskeskuksen hallituksen nimittämä TOT-johtokunta. Tutkintaryhmien käytännön toimintaa johtavat Tapaturmavakuutuskeskuksen työturvallisuusasiantuntijat. Tutkinnan tarkoitus on työtapa-
turmien ja ammattitautien torjuminen sekä näiden tutkintojen tehokas työstäminen. Tutkinnassa selvitetään sattuneiden vahinkojen syyt ja vahinkoja edeltäneet tapahtumat. Samassa yhteydessä pyritään laatimaan suosituksia työturvallisuuden kehittämiseksi. Kehittämissuosituks^{et} voivat koostua työympäristöön ja ihmisten toimintaan liittyvistä välittömistä tekijöistä sekä organisaatioiden toimintaan liittyvistä välillisistä tekijöistä. Tutkintaan ottamisen perusteiden kannalta on erityisen merkityksellistä, voiko tutkinta tuottaa tärkeää lisätietoa vastaavien työtapa-
turmien ja vaaratilanteiden ennaltaehkäisemiseksi. (Tapaturmavakuutuskeskus, 2022b)

2.3 Työtapaturmat Suomessa

Tapaturmavakuutuskeskus kerää työtapa-
turmien ja ammattitautien tilastoja vakuutusyhtiöiltä ja tuottaa analyyseja kerätyn tiedon pohjalta. Työturvallisuus on osa työoikeuden kokonaisuutta. 1970-luvulla otettiin työsuojelu-termi käyttöön ja se onkin nykyään vakiintunut termi. Nykyään käytetään lisäksi termiä työympäristö, jolla kuvataan toiminnan sekä sääntelyn kohdetta. (Saloheimo, 2016)

Työympäristön riskiä on saatu vähennettyä huomattavasti uusilla suojeluvaatimuksilla. Näistä esimerkkeinä mainittakoon asbestille altistuminen, jota on saatu vähennettyä huomattavasti pakollisten turvatoimien myötä. Tapaturma- ja ammattitautien tilastoihin vaikuttavat normien ja valvonnan lisäksi monet muutkin asiat, kuten muutokset elinkeinorakenteessa, tuotantomenetelmien muutokset, taloudelliset suhdanteet, työllisyystilanne sekä tuotannon käyntiaste. Työtapa-
turmien riski vaihtelee myös suuresti toimialan mukaan. Yksi vaarallisimmista aloista on rakentaminen. Suhdannevaihteluihin herkästi reagoivilla toimialoilla on yleisesti myös korkeat tapaturmataajuustasot. Tapaturmataajuudella tarkoitetaan tehtyjen työtuntien ja sattuneiden tapaturmien suhdetta. Kokonaisuudessa työtapa-
turmien määrä seuraa eritoten rakentamisen, teollisuuden sekä kuljetuksen toimialojen työtuntien määrän muutosta. (Saloheimo, 2016)

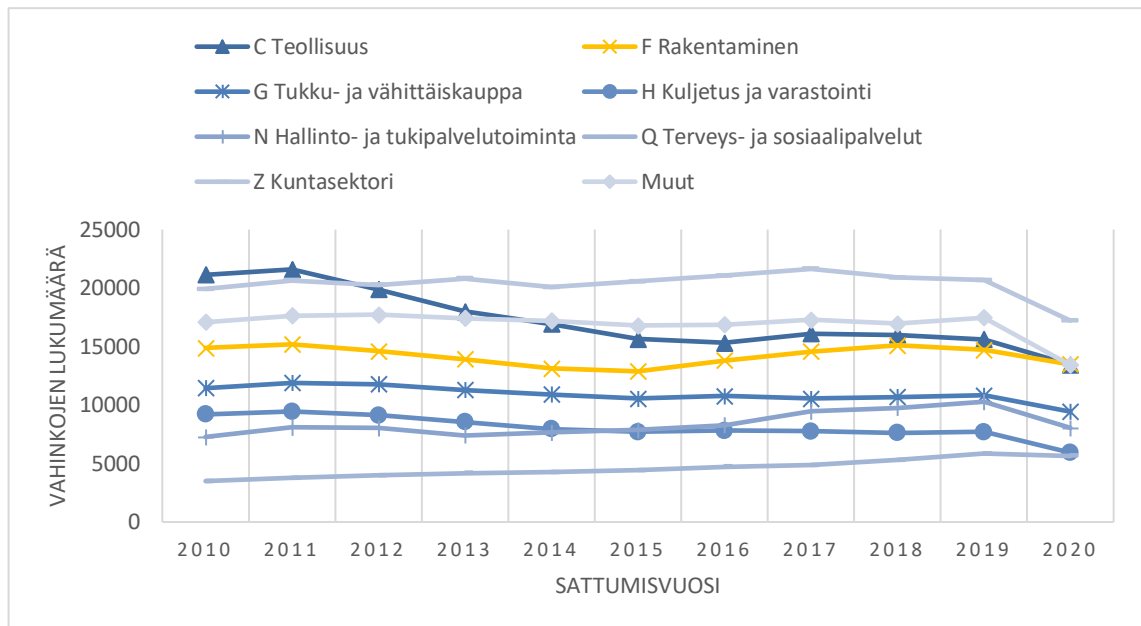
Työtapa-
turmia sattuu Suomessa vuosittain yli 100 000 kappaletta Kyseinen luku sisältää sekä palkansaajien työpaikkatapaturmat että työmatkatapaturmat. Vuosittain työpaikkatapaturmien määrä on tavallisesti ollut yli 100 000 kappaletta kun taas tuoreimman saatavissa olevan tiedon mukaan niitä on ollut hieman yli 90 000 kappaletta. Ko-

ronasta johtuva poikkeustilanne on vaikuttanut alentavasti työtapaturmien määriin. Työtapaturmat jaetaan palkansaajille sekä yrittäjille sattuneisiin työtapaturmiin. Valtaosa työtapaturmista sattuu palkansaajille, sillä yrittäjien osuus työtapaturmista on vain noin viisi prosenttia. (Tapaturmavakuutuskeskus, 2020)

Yleisimmin työtapaturmat aiheuttavat vahingon kärsineelle vain lyhytkestoista työkyvyttömyyttä, mutta voi olla, ettei tapaturma vaikuta työkykyyn lainkaan. Vuoden 2020 arvioiden mukaan palkansaajien työtapaturmista aiheutui korkeintaan kolmen päivän mittaisen työkyvyttömyys noin 70 %:lle vahingon kärsineistä. Lähes 24 %:lle vahingon kärsineistä aiheutui 4–30 päivän mittainen työkyvyttömyys. Palkansaajien kuolemaan johtaneita työpaikkatapaturmia on sattunut viimeisten viiden vuoden aikana 14–27 kappaletta vuodessa. (Tapaturmavakuutuskeskus, 2020)

2.3.1 Työtapaturmat rakennusteollisuudessa

Palkansaajien työpaikkatapaturmien määrät ovat olleet laskusuunnassa rakentamisen päätoimialalla vuodesta 2018 lähtien, mutta kehitys vaikuttaisi olevan vuoden 2021 tiedon mukaan olevan tasaantumassa (kuva 1). Kuvan 1 kuvaajien tarkat lähtötiedot löytyvät liitteestä A. Mainittakoon, että suuri osa rakennustyömailla sattuneista työtapaturmista sattuu muille kuin rakentamisen alan palkansaajille. Tällaisia ovat muun muassa aliurakoitsijat, asentajat, vuokratyöntekijät sekä tavarantoimittajat. Vuonna 2021 palkansaajille sattui yhteensä 13445 työpaikkatapaturmaa. Työpaikkatapaturmat jaotuitivat alaluokkiin seuraavasti: talonrakentaminen 5042 kappaletta, maa- ja vesirakentaminen 968 kappaletta ja erikoistunut rakennustoiminta 7434 kappaletta. (Tapaturmavakuutuskeskus, 2020; Tapaturmavakuutuskeskus, 2022a)

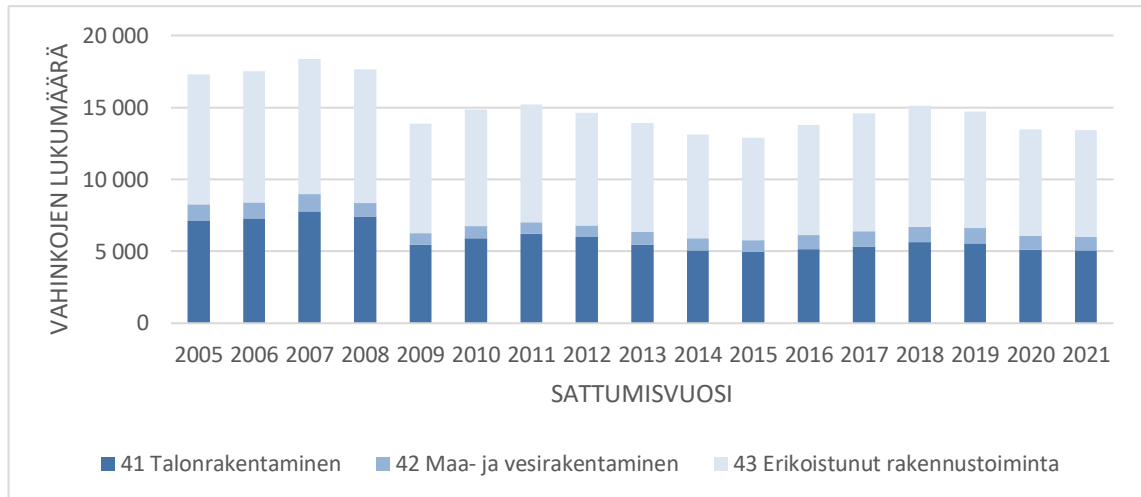


Kuva 1. *Palkansaajien työpaikkatapaturmat päätoimialan mukaan sattumisvuosittain*

Rakentamisen päätoimiala jaetaan Tilastokeskuksen vuoden 2008 toimialaluokituksen mukaan kolmeen alaluokkaan:

- talonrakentaminen
- maa- ja vesirakentaminen
- erikoistunut rakennustoiminta.

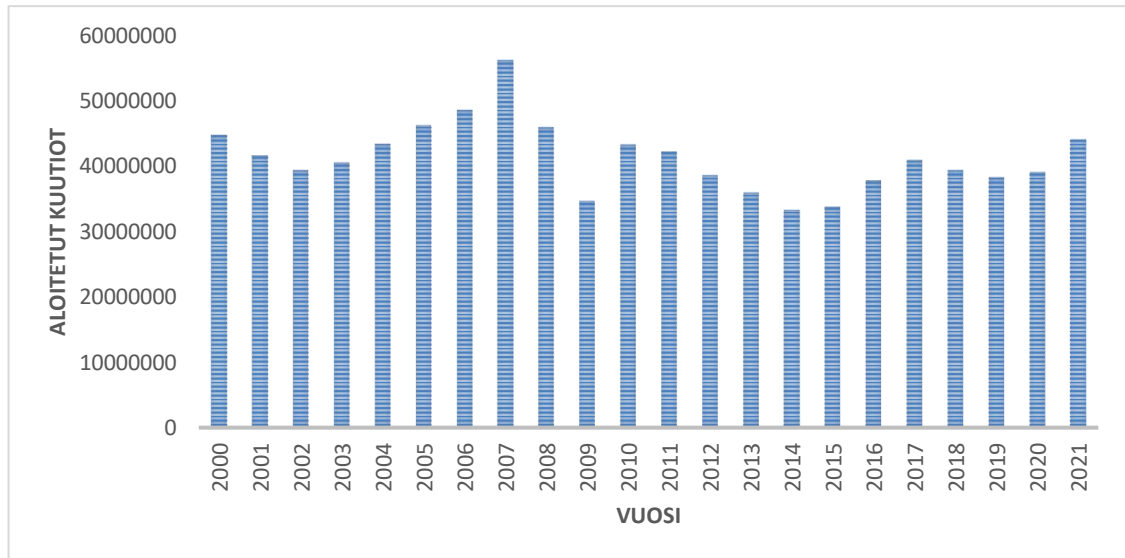
Rakentamisen päätoimialalla suurin osa vuosittaisista työpaikkatapaturmista kohdistuu erikoistuneeseen rakennustoimintaan, mutta myös talonrakentamiseen kohdistuu paljon työpaikkatapaturmia. Maa- ja vesirakentamisessa sattuu vain vähäinen määrä työpaikkatapaturmia vuosittain (kuva 2).



Kuva 2. *Palkansaajien työpaikkatapaturmat sattumisvuosittain rakentamisen päätoimialalla, jaoteltuna talonrakentamiseen, maa- ja vesirakentamiseen ja erikoistuneeseen rakennustoimintaan*

Rakennusteollisuudessa on jo pidempään sattunut suhteellisesti paljon vakavia loukkaantumisia ja kuolemantapauksia. Rakennusteollisuus on siis yksi vaarallisimmista aloista maailmassa. Rakennusteollisuutta on kuvattu monimutkaiseksi, dynaamiseksi ja ainutlaatuiseksi alaksi, johon monesti liittyvät riskialttiit työtavat. Muutosta ei ole saatu aikaiseksi, joten seuraavana askeleena teollisuus ja akateeminen maailma ovat alkaneet tehdä yhteistyötä siirryttäessä säädösten mukaisesta toimintatavasta enemmän proaktiiviseen ja ennustavaan lähestymistapaan. Turvallisuuden ennustaminen sisältää poikkeuksetta menneen tiedon hyödyntämistä tulevan ennakoimisessa. Tämäkin tutkimus pyrkii osaltaan vastaamaan tähän tarpeeseen ennustavien menetelmien avulla. (Bhagwat & Delhi, 2021; Bhagwat et al., 2022; Hallowell et al., 2020)

Ennustetusta vahinkomäärästä saadaan tarkempi tulos lisäämällä malliin selittävä tekijä. Tässä tutkimuksessa selittävä tekijänä käytetään rakennushankkeiden aloitettua tilavuutta kuutioina, koska se kuvaa hyvin rakennettavan rakennuskannan määrää. Aiemmissa työtapaturmien ennustemalleissa on käytetty selittävä tekijänä muun muassa liikevaihtoa, mutta keskimääräistä korkeamman inflaation oloissa liikevaihto saattaa lisätä ennusteharhaa. Rakennusteollisuus ilmoittaa tiedot rakennushankkeiden aloitetuista kuutioista Tilastokeskukselle, joten tiedot kuutioista tähän tutkimukseen on haettu Tilastokeskuksen Statfinn-tietokannasta. Vuosittain rakennushankkeiden aloitetuissa kuutioissa on paljonkin vaihtelua. Volyymiin vaikuttavat esimerkiksi talouden tunnusluvut aina inflaatiosta korkoihin (kuva 3).



Kuva 3. Rakennusluvanvaraisen uudisrakentamisen volyymi aloitettuina kuutioina (Lähde: Tilastokeskus, Rakennus- ja asuntotuotanto)

3. ENNUSTEMENETELMÄT

Analytiikasta yleisesti puhuttaessa tarkoitetaan nykyään muun muassa tilastotieteen, ennustamisen, koneoppimisen, tekoälyn ja syväoppimisen menetelmiä. Mitä nämä kaikki sitten tarkoittavat niin siitä on montaa näkemystä. Käsite robotiikkakin on otettu käyttöön, Robotic Process Automation (RPA) eli suomennettuna ohjelmistorobotiikka kuulostaa hienolta, vaikka kyse on niinkin arkisesta asiasta kuin prosessien automatisoinnista. Prosessien automatisointia on ollut käytössä jo aiemmin esimerkiksi käyttöliittymätason automatiikkana ja ensimmäiset askeleet ovat olleet testausautomaatiikan hyödyntämistä sovelluksille. Tätä voidaan hyödyntää muun muassa korvausjärjestelmän testaamisessa.

Tässä tutkielmassa käsitellään syväoppivia neuroverkkoja ja niiden sovelluksia sekä esitellään lyhyesti nykyisin organisaatioissa käytetty VAR-malli. Työssä esitetään neuroverkkojen perusrakenne sekä tarkemmin LSTM- ja GRU-neuroverkkorakenteet. Lisäksi tutustutaan tarvittavilta osin koneoppimisen perusteisiin. Lopuksi sovelletaan edellä mainittuja neuroverkkorakenteita työtapaturmavahinkojen lyhyen aikavälin ennustamiseen.

Tilastollisen mallin avulla pyritään kuvaamaan mahdollisimman hyvin mallin kohteena olevan vastemuuttujan sekä mallia selittävien muuttujien välistä yhteyttä. Rakennetut mallit sovitetaan käytettävään aineistoon ja tietyin kriteerein niistä valitaan käyttötarkoitukseen sopivin malli. Tutkimusaineistona käytetään Tapaturmavakuutuskeskuksen keräämää aineistoa vuodesta 1995 vuoteen 2021. Aineisto on rajattu tarkemmin rakennusteollisuuden päätoimialan alaluokkiin talonrakentaminen sekä erikoistunut rakennustoiminta. Aineiston tulee olla riittävän laaja käytettävissä oleviin menetelmiin liittyen. Mallin selittävänä muuttujana käytetään Tilastokeskuksen tilastoimia aloitettujen rakennushankkeiden määriä kuutioissa mitattuina.

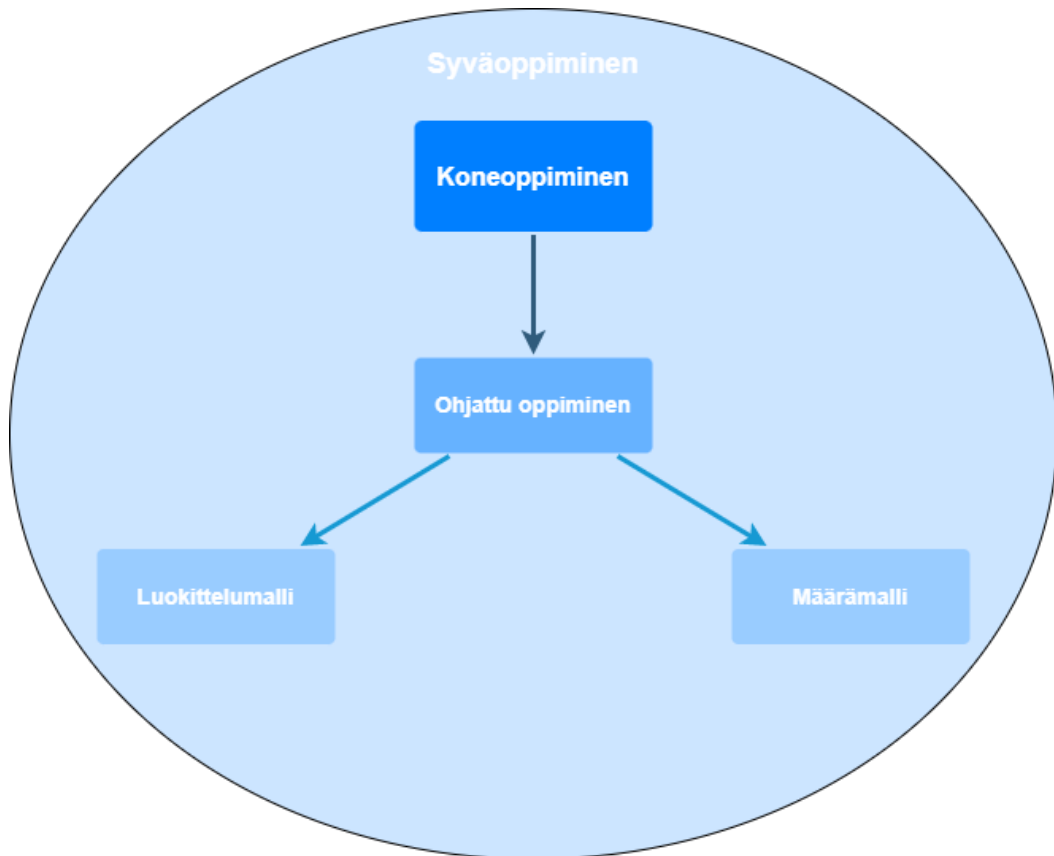
3.1 Vektoriautoregressiivinen malli (VAR-malli)

VAR-malli on moniyhtälöaikasarjamalli, jossa jokaista selitettävää muuttujaa selittävät omat sekä muiden selitettävien muuttujien viipeet. VAR-malli ennustaa hyvin, jos muuttujat ovat vahvasti toisistaan riippuvia ja lisäksi ajanhetken t arvoja voidaan selittää

edeltävien ajanhetken $t-1$ arvoilla. VAR-mallit sopivat ennustamiseen paremmin kuin esimerkiksi yleinen lineaarinen tilastollinen malli, koska endogeeniset ja selittävät muuttujat ovat interaktiossa keskenään. Toisin sanoen jokaista selitettävää muuttujaa selitetään omien viipeiden lisäksi muiden selitettävien muuttujien viipeillä. VAR-mallien kyky ennustaa on todettu hyväksi, sillä muuttujien väliset korrelaatiot voidaan spesifioida joustavasti ja yleisesti kaikkina ajanhetkinä. Organisaatiossa aiemmin käytetty malli ennustaa kaikkia työtapaturmia, joten sitä ei ole mielekästä verrata tässä tutkielmassa rakennettavaan neuroverkkomalliin. (Pecican, 2010; Suomen Pankki, 1989)

3.2 Koneoppiminen

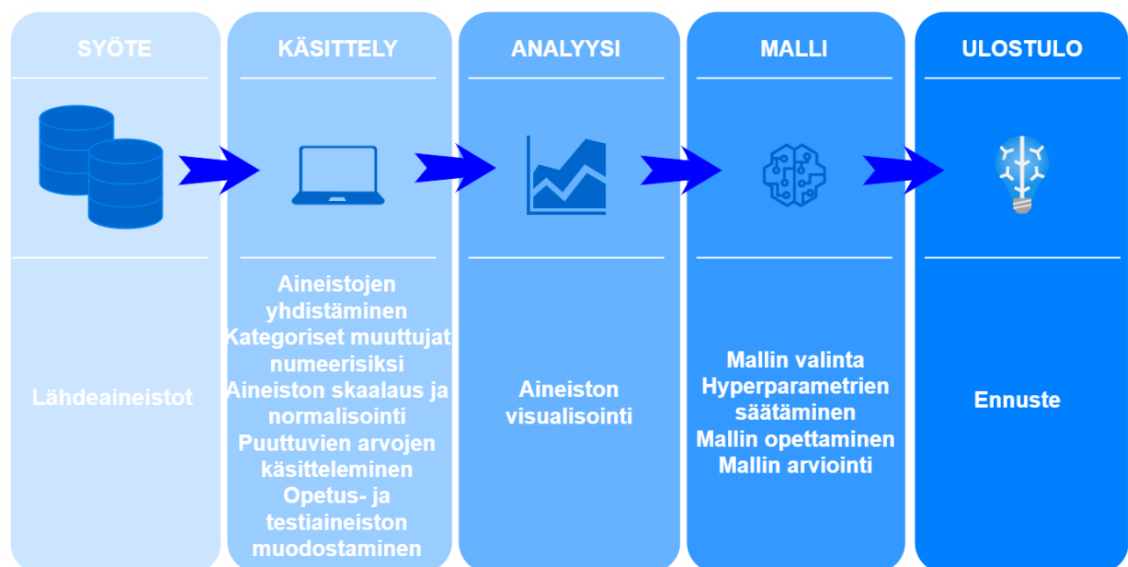
Koneoppimisen tavoitteena on pyrkiä rakentamaan parhaiten sopiva malli kuvaamaan vastemuuttujan ja selittävien muuttujien välistä yhteyttä, joten ajatus on hyvin sama mitä edellä viitatussa tilastollisessa mallinnuksessa. Koneoppimisen (Machine Learning) tekniikan periaatteet on kehitetty jo 1950- ja 1960-luvulla. Koneoppiminen käyttää algoritmeja, jotka taas tarvitsevat dataa. Algoritmien suorituskyky tai tulokset eivät parane datan määrän kasvaessa. Syvien neuroverkkojen (Deep Learning) tapauksessa taas tulokset paranevat datan määrän kasvaessa. Koneoppimisen ennustemallit jaetaan pääsääntöisesti kahteen pääluokkaan: luokittelumalleihin ja määrämalleihin (kuva 4). Luokittelumallien avulla ennustetaan tietyn ennalta määritellyn tapahtuman todennäköisyyttä. Ideana on vastata kysymykseen, tapahtuuko jokin asia vai ei. Malliin voidaan sisällyttää myös todennäköisyyttä jollekin asialle tapahtuvaksi. Määrämallien avulla ennustetaan jonkin asian suuruusluokkaa. Luokittelumallissa käytettäviä tekniikoita nimitetään klassifikaatiomalleiksi, kun taas määrämallien tekniikoita nimitetään regressiomalleiksi. (Kananen, 2019)



Kuva 4. Koneoppimisen ennustemallien pääluokat

Koneoppimisen ennustemallien avulla pyritään täydentämään tai korvaamaan asiantuntijoiden tekemiä arvioita sekä manuaalista päätöksentekoa. Ennustemallien avulla saadaan tuloksia suuresta määrästä dataa nopeasti ja edullisesti. Ennustemallien tuottamien tulosten tarkkuutta tulee arvioida tapauskohtaisesti. Ennustemalli perustuu sille syötettyyn aineistoon ja se toimii tilastollisten jakaumien mukaan. Koneoppiminen voidaan havainnollistaa myös Loyn esittelemän työkulkukaavion avulla (kuva 5). Kaaviossa aineisto toimii syötteenä ja lopputuloksena haluttu ulostulo. Aineisto toimii aina syötteenä. Aineisto voi tulla useammasta lähteestä ja se voi olla erilaisissa muodoissa. Riippuen koneoppimisen mallista aineisto voi olla esimerkiksi kuva tai rivipohjaista aineistoa. Seuraavassa vaiheessa on aineiston käsittely. Aineisto tulee yleensä muokata sellaiseen muotoon, että sitä voidaan hyödyntää käytettävässä mallissa. Koneoppimisen mallit vaativat opetukseen numeerista aineistoa. Jos lähtöaineistossa on kategorisia muuttujia, niin ne tulee muuttaa numeerisiksi muuttujiksi. Lähtöaineisto voi olla puutteellista, joten sille voi olla tarpeen tehdä skaalausta ja normalisointia. Puutteellisen aineiston käsittelemiseen voidaan käyttää useampia keinoja, joista helpoin on poistaa ne havainnot tai muuttujat, jotka eivät täytä vaadittuja kriteereitä. Aineiston käsittelyn lopuksi aineisto jaetaan opetus- ja testausaineistoon. Jako tehdään siitä syystä, että

koneoppimisen mallia opetetaan ja sovitetaan ainoastaan opetusaineistolla. Testausaineistolla taasen arvioidaan mallin soveltuvuutta tilanteeseen. Seuraavaksi aineistolla suoritetaan kuvainnollista analyysiä visualisoinnin avulla. Tällä vaiheella lisätään informaatiota malliin saatujen havaintojen pohjalta. Seuraavassa vaiheessa tapahtuu mallin rakentaminen. Mallin rakentamisen vaiheessa hyödynnetään valittuja koneoppimisen algoritmeja. Opettaminen tapahtuu opetus esimerkkien avulla ja samalla muun muassa hyperparametreja säädetään opetusaineistolla ja testataan testiaineistolla. Tarkoitus on, että koneoppimisen malli kykenee yleistämään oppimansa ja toimii lopulta riittävällä tarkkuudella tuntemattomalle aineistolla. Lopuksi voidaan arvioida mallia testiaineiston avulla. (Kananen, 2019; Loy, 2019)

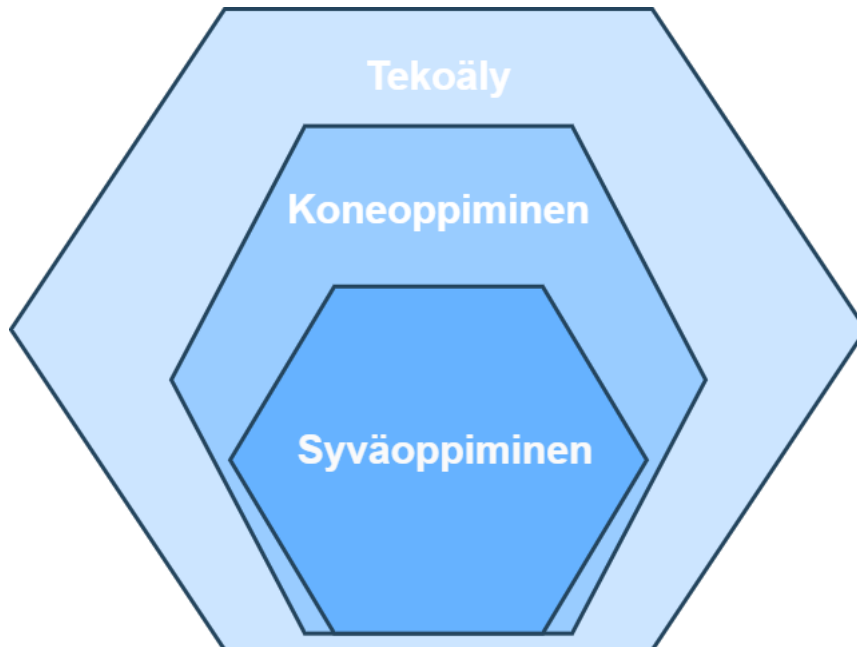


Kuva 5. Koneoppimisen työskulkukaavio (mukaillen Loy 2019)

3.3 Neuroverkot

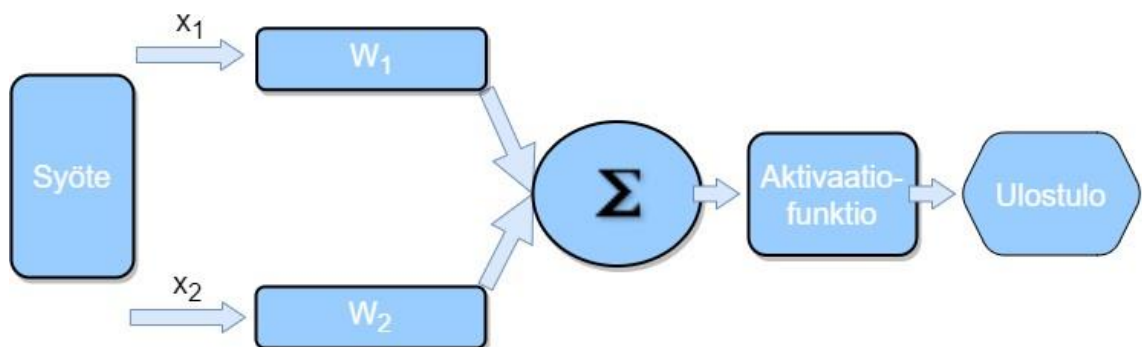
Tekoäly tulee englanninkielisestä termistä Artificial Intelligence (AI). Suomen kielessä esiintyy myös termit keinoäly, koneäly, koneoppiminen ja syväoppiminen. Usein keskustelussa nämä termit menevät sekaisin. Tekoälyn osajoukkoja on esitelty kuvassa 6. Tekoälyn tekniikka perustuu ohjelmointiin, matematiikkaan sekä tilastotieteeseen. Koneoppimisen tavoitteena on pyrkiä rakentamaan parhaiten sopiva mallia kuvaamaan vastemuuttujan ja selittävien muuttujien välistä yhteyttä, mihin jo viitattiin aiemmin. Syväoppivat (Deep Learning) neuroverkot ovat kehittyneempi versio neuroverkkoarkitehtuurista. Syväoppivat verkot ovat kehittyneet paljon viime aikoina ja se on lisännyt kiinnostusta tekoälyn hyödyntämisestä kohtaan. Syväoppivaan verkkorakenteeseen on lisätty kognitiivisia teknologioita, kuten oppimista, muistia ja logiikkaa. Neuroverkoissa on paljon muuttujia ja ne ovat moniulotteisia. Syväoppivat neuroverkot ovat kehittyneet

paljon viimeisen vuosikymmenen aikana, koska dataa ja laskentatehoa on paljon saatavilla. Neuroverkkojen avulla pystytään tehokkaasti tarkastelemaan epälineaarisia ilmiöitä, sillä perinteisin keinoin toteutettuna matemaattiseen yhtälöön kertyy nopeasti valtava määrä termejä, mikä taas vaikeuttaa laskentaa. (Kananen, 2019)



Kuva 6. *Tekoälyn osajoukot*

Neuroverkot voidaan esittää yksinkertaisesti matemaattisella funktiolla, joka kuvaa annetun syötteen haluttuna ulostulona. Asian selventämiseksi tarkastellaan yksikerroksista neuroverkkoa, mitä kutsutaan perseptroniksi. Perseptroni voidaan esitellä oheisen kuvan avulla (kuva 7):



Kuva 7. *Perseptronin rakenne (mukaillen Loy 2019)*

Peräkkäiset matemaattiset funktiot (x), jotka kytkeytyvät toisiinsa muodostavat neuroverkon. Painokertoimien arvo (w) muodostaa neuroverkkoon niin sanotun opitun tiedon. Painokerroin kuvastaa, miten voimakas neuronien välinen yhteys on. Painoker-

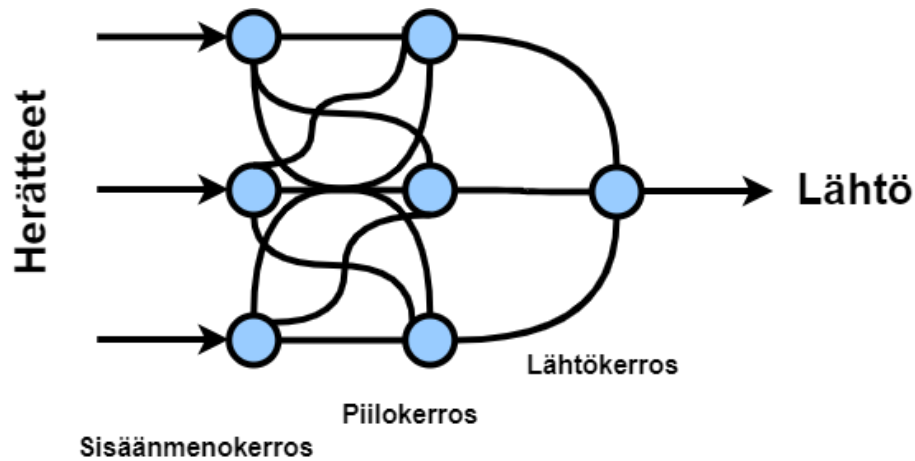
toimia kutsutaan myös synaptisiksi painoiksi. Painokertoimet ja sisäänsyötön kertolaskutoimitusten tulokset summataan yhteen (Σ) ja sen jälkeen ne ajetaan aktivaatiofunktion läpi ennen ulostuloa. Aktivaatiofunktio on käytännössä neuronin funktio. Aktivaatiofunktion avulla neuroverkoissa voidaan hyödyntää epälineaarisen mallin mukanaan tuomia hyötyjä. Aktivaatiofunktio rajoittaa seuraavan kerroksen sisäänmenoa, jotta summa ei kasva äärettömäksi verkon alimmissa kerroksissa. Suosituimpia aktivaatiofunktioita ovat Sigmoid-, Tanh- ja ReLU-aktivaatiofunktiot. Lista aktivaatiofunktioista on esitetty liitteessä B. Oppimisalgoritmeilla verkon lukuisia painoarvoja muokataan niin, että virhe alkuperäisen ja verkon tuottaman ulostulon välillä on mahdollisimman pieni. Neuroverkot koostuvat siis toisiinsa kytketyistä peräkkäisistä matemaattisista funktioista ja neuroverkon älykkyys perustuu painokertoimien arvoon ja määrään. (Kananen, 2019)

Neuroverkkojen osalta datassa olevia muuttujia nimitetään herätteiksi. Neuroverkkojen keskeinen osa on neuroni (kuva 8), jolla voi olla monta herätettä (input), mutta vain yksi lähtöarvo (output). Neuronin toiminta perustuu aktivaatiofunktioon, jonka lähtöarvo riippuu herätteen tasosta. (Kananen, 2019)

Neuroverkko koostuu yksinkertaisesti kolmesta erityyppisestä kerroksesta (kuva 8):

- sisäänmenokerros (Input Layer)
- piilokerros (Hidden Layer)
- lähtökerros (Output Layer). (Kananen, 2019)

Sisäänmenokerroksessa ei tapahdu varsinaisesti laskentaa, vaan neuroverkko vastaanottaa datan ja välittää sen eteenpäin neuroverkossa. Piilokerroksessa tapahtuu varsinainen neuroverkon prosessointi ja painokertoimien määrittäminen. Laskentaprosessi tapahtuu koulutusvaiheessa niin, että painokertoimet saattavat olla aluksi todella huonoja, joten neuroverkko tuottaa ensimmäisillä laskentakierroksilla epätarkkoja tuloksia. Painokertoimien päivittäminen ja sitä kautta tulosten tarkentuminen tapahtuu vertaamalla neuroverkon antamaa ulostuloa datan tiedettyyn arvoon. Lähtökerros tuottaa lopullisen lähdön, joka perustuu edeltävän kerroksen tuloksiin. (Kananen, 2019)



Kuva 8. Yksinkertaisen neuroverkon esityskaavio (mukaillen Kananen 2019)

Takaisinkytketyt neuroverkot (Recurrent Neural Network - RNN) ovat verkkoja, jotka säilyttävän informaation. Niitä käytetään esimerkiksi puheentunnistukseen. Takaisinkytkettyjen neuroverkkojen heikkous on kuitenkin lyhytkestoinen muisti. Jos sekvenssi on riittävän pitkä niin informaation kulkeutuminen aiemmista ajanhetkistä viimeisiin voi muodostua ongelmaksi (Vanishing Gradient Problem).

On olemassa erilaisia neuroverkkoja: autoenkoodaajat, konvoluutioneuroverkko, eteenpäin kytketty neuroverkko ja takaisin kytketty neuroverkko. Tässä työssä tarkastellaan pitkä lyhytkesto -muistiverkkoa (Long Short-Term Memory - LSTM) sekä avainnetuksi toistuvaksi soluksi nimettyä neuroverkkoa (Gated Recurrent Unit - GRU).

3.3.1 LSTM-arkkitehtuuri

LSTM-arkkitehtuurin ovat esitelleet alun perin vuonna 1997 Sepp Hochreiter ja Jürgen Schmidhuber. LSTM-neuroverkot perustuvat RNN-arkkitehtuuriin, joissa on mahdollista oppia pitkien aikajaksojen sekvensseistä. LSTM-arkkitehtuurit ovat syväoppimisen monimutkainen osa-alue ja se on luotu välttämään pitkän ajanjakson riippuvuusongelmia. LSTM-kerroksen laskentaprosessi on suhteellisen monimutkainen. Sillä on pidempi muistikapasiteetti ja se suoriutuu perinteistä takaisinkytkettyä neuroverkkoa (RNN) paremmin suurimmassa osassa sekvenssitehtävistä. Sekvenssimallit ovat koneoppimisen malleja, jotka käyttävät syötteenä tai ulostulona aineiston sekvenssejä. Sekvenssiaineisto voi sisältää muun muassa tekstiä tai aikasarjaa. LSTM-arkkitehtuurissa piilokerros koostuu takaisinkytketyistä aliverkoista, joita kutsutaan muistilohkoiksi. Jokainen lohko sisältää joukon sisäisiä yksiköitä tai soluja, joiden aktivointia kontrolloidaan kolmella portilla: sisäänmenoportti (input gate), unohtusportti (forget gate) ja lähtöportti

(output gate). LSTM-verkossa on muistilohko, joka sisältää yhden tai useamman muistisolun ja parin adaptiivisia, epälineaarisia avainnyksiköitä (Gating Units), jotka ohjaavat herätettä ja lähtöä lohkon jokaiseen soluun. Jokaisessa muistisolussa on itsekytkettyvä lineaarinen vakiovirheen karuselli (Constant Error Carousel – CEC), jonka aktivointia kutsutaan solun tilaksi. Vakiovirheen karuselli ratkaisee aiemmin esitetyn takaisinkytkettyjen neuroverkkojen (RNN) ongelman lyhytkestoisien muistin suhteen. (X. Liu et al., 2021; Graves et al., 2009; Gers et al., 2000)

LSTM-arkkitehtuuri koostuu useista solmuista, jotka on kytketty toisiinsa arkkitehtuurin mukaisesti. Jokaisessa solmussa on neljä toimintoa: muistin sisäänkirjoitus, unohtaminen, muistin lukeminen ja ulostulo. Näiden toimintojen avulla LSTM-malli pystyy säilyttämään tietoa pitkällä aikavälillä ja käyttämään sitä ennustettaessa. LSTM-arkkitehtuuri on yksi tärkeimmistä keksinnöistä syväoppimisalalla, ja se on auttanut kehittämään monia tehokkaita ennustusmalleja eri sovelluskohteisiin. (W. Liu et al., 2020)

LSTM on mallinnusmenetelmä, jota käytetään tekstin ja muun aikasarjatiedon ennustamisessa. LSTM-mallit ovat yksi tyyppi syväoppimismenetelmistä, ja ne eroavat tavallisista syväoppimismalleista siinä, että ne pystyvät muistamaan tietoa pitkällä aikavälillä ja käyttämään sitä ennustettaessa. Tämä tekee LSTM-malleista erityisen hyödyllisiä esimerkiksi tekstikäsittelyssä, kielentunnistuksessa ja muissa sovelluksissa, joissa tarvitaan kykyä ymmärtää tekstiä ja sen merkitystä pitkällä aikavälillä. LSTM-arkkitehtuurin uniikit ominaisuudet ovat laajentaneet niiden käyttöä koneoppimisen saralla. Sovelluskohteina on muun muassa aineiston luokittelu sekä puheen ja kirjoittamisen tunnistaminen. Viimeisimpänä LSTM-malleja on käytetty myös dynaamisten prosessien, kuten lämpötilan ohjaamiseen ja mallit ovat erittäin tarkkoja testiaineistoon verrattuna. (Zarzycki & Ławryńczuk, 2021; Zazo et al., 2016)

3.3.2 GRU-arkkitehtuuri

GRU-arkkitehtuuri on esitelty ensi kertaa vuonna 2014 lyhytkestoisien muistin ongelman ratkaisemiseksi takaisinkytkettyjen neuroverkkojen tapauksessa. Tätä muunnelmaa kutsutaan avainnetuksi toistuvaksi soluksi (Gated Recurrent Unit – GRU). (Cho et al., 2014). GRU-arkkitehtuuri ratkaisee LSTM-arkkitehtuurin tapaan aiemmin esitetyn takaisinkytkettyjen neuroverkkojen (RNN) ongelman. LSTM-arkkitehtuurin solu, joka koostui kolmesta portista, rakentuu GRU-arkkitehtuurissa vain kahden portin solusta: päivitysportista (update gate) sekä resetointiportista (reset gate). (X. Liu et al., 2021)

LSTM-neuroverkko ei ole taipuvainen ongelmaan informaation kulkeutumisesta aiemmista ajanhetkistä viimeisiin (Vanishing Gradient Problem), mutta LSTM-arkkitehtuurin rakenne on verrattain monimutkainen, laskentakustannus on korkea ja parametrit suuria. Tästä johtuen LSTM-arkkitehtuurin laskentaprosessia on yritetty yksinkertaistaa vähentämällä etenkin porttien määrää. Tutkimukset ovat osoittaneet, että unohdusportti on kaikkein tärkein porttikäsittely LSTM-arkkitehtuurissa ja on osoitettu, että yksinkertaistettu versio verkosta vain yhdellä unohdusportilla on parempi, kuin perinteinen LSTM-arkkitehtuuri. Näin on muodostunut käyttöön laajalti käytetty takaisinkytkettyjen verkkojen (RNN) variantti, LSTM-arkkitehtuurista yksinkertaistettu GRU-arkkitehtuuri. (Long, 2022)

GRU-malli perustuu LSTM-malliin, mutta siinä on tiiviimpi ja yksinkertaisempi rakenne. Malli ratkaisee lyhytkestaisen muistin ongelman edellä mainittujen kahden portin avulla. Portit koostuvat kahdesta vektorista, jotka manipuloivat verkon tai kerroksen tietoa, joka virtaa haluttuun lähtöön. Nämä kaksi porttia voidaan opettaa säilyttämään pitkäkestäisen muistin tiedon ilman, että oleellista tietoa jää puuttumaan ennusteesta. Yksilöllisen piilosolun avulla voidaan saada tietoa riippuvuuksista, jotka määräytyvät vastaavien avainmekanismien toimintataajuuden perusteella. Päivitysportti ratkaisee, kuinka paljon informaatiota aiemmista ajanhetkistä tarvitaan tuleviin ajanhetkiin ja samaan aikaan resetoitiportti avustaa mallia määrittämään, kuinka paljon mennyttä informaatiota tarvitsee ottaa käyttöön. Tämän rakenteen avulla GRU-malli voi hyödyntää tarvittavan menneisyyden informaation ja sen avulla vähentää huomattavasti lyhytkestäisen muistin ongelmaa. GRU-mallissa ei siis ole lisätty muistisolua informaation säilyttämiseen, joten mallissa hallinnoidaan informaatiota solun sisällä. (Shahid et al., 2020; . X. Liu et al., 2021)

GRU-arkkitehtuurin erityisominaisuudet ovat lisänneet sen käyttöä. GRU-arkkitehtuuria käytetään usein esimerkiksi kielenkäännöksissä, tekstiluokituksessa ja kielimallinnuksessa, koska ne pystyvät tehokkaasti havaitsemaan jatkuvan datan rakenteen ja riippuvuudet. Kyseistä arkkitehtuuria sovelletaan muun muassa aloilla, joissa on tärkeää mallintaa pitkäaikaisia riippuvuuksia, kuten taloudellisessa ennustamisessa ja puheentunnistuksessa. Lisäksi arkkitehtuuria käytetään kasvojen ilmeiden tunnistuksessa, tuotteiden valmistusvirheiden havaitsemisessa sekä vihapuheen tunnistamisessa etenkin sosiaalisen median alustoilla. (Zarzycki & Ławryńczuk, 2021)

4. TYÖTAPATURMIEN MÄÄRÄN ENNUSTAMINEN

Tässä työssä käytetään kahta aikasarjaa ennustamaan työpaikkatapaturmien lyhyen aikavälin (12 kuukautta) lukumäärää organisaation työturvallisuusyksikössä. Ensimmäinen aikasarja on Tapaturmavakuutuskeskuksen keräämiä työtapaturmavakuutuksen tilastoja vakuutusyhtiöiltä. Toinen aikasarja on Tilastokeskuksen keräämiä aloitettuja rakennuskuutioita rakennus- ja asuntotuotannossa.

4.1 Ongelman asettelu

Tutkimuksen tarkoituksena on verrata organisaatiossa nykyisin käytettävää VAR-ennustemallia tässä työssä käytettäviin LSTM- ja GRU-malleihin. Tapaturmavakuutuskeskuksessa on käytössä vahinkojen ennustamiseen aiemmin toteutettu VAR-malli. VAR-malli on lineaarinen ja neuroverkkojen avulla voidaan hyödyntää epälineaarisen mallin mukanaan tuomia hyötyjä.

4.2 Aineisto

Tapaturmavakuutuskeskus kerää työtapaturmavakuutuksen tilastoja vakuutusyhtiöiltä. Aineistossa on mukana kaikki vakuutusyhtiöille ilmoitetut palkansaajien työpaikkatapaturmat rakentamisen päätoimialan tietyissä alaluokissa. Rakentamisen päätoimiala jaetaan Tilastokeskuksen vuoden 2008 toimialaluokituksen mukaan yhteensä kolmeen alaluokkaan:

- talonrakentaminen
- maa- ja vesirakentaminen
- erikoistunut rakennustoiminta.

Aineisto sisältää rakentamisen päätoimialan talonrakentaminen sekä erikoistunut rakennustoiminta alaluokkien vahinkomäärät kuukausittain sattumisvuosina 1995–2021. Talonrakentamisen ja erikoistuneen rakennustoiminnan alaluokat kattavat vuosittain n. 93 % rakennusalan palkansaajille sattuneista työtapaturmista. Aineistossa ei ole mukana työmatkalla sattuneita työtapaturmia eikä yrittäjille sattuneita työtapaturmia.

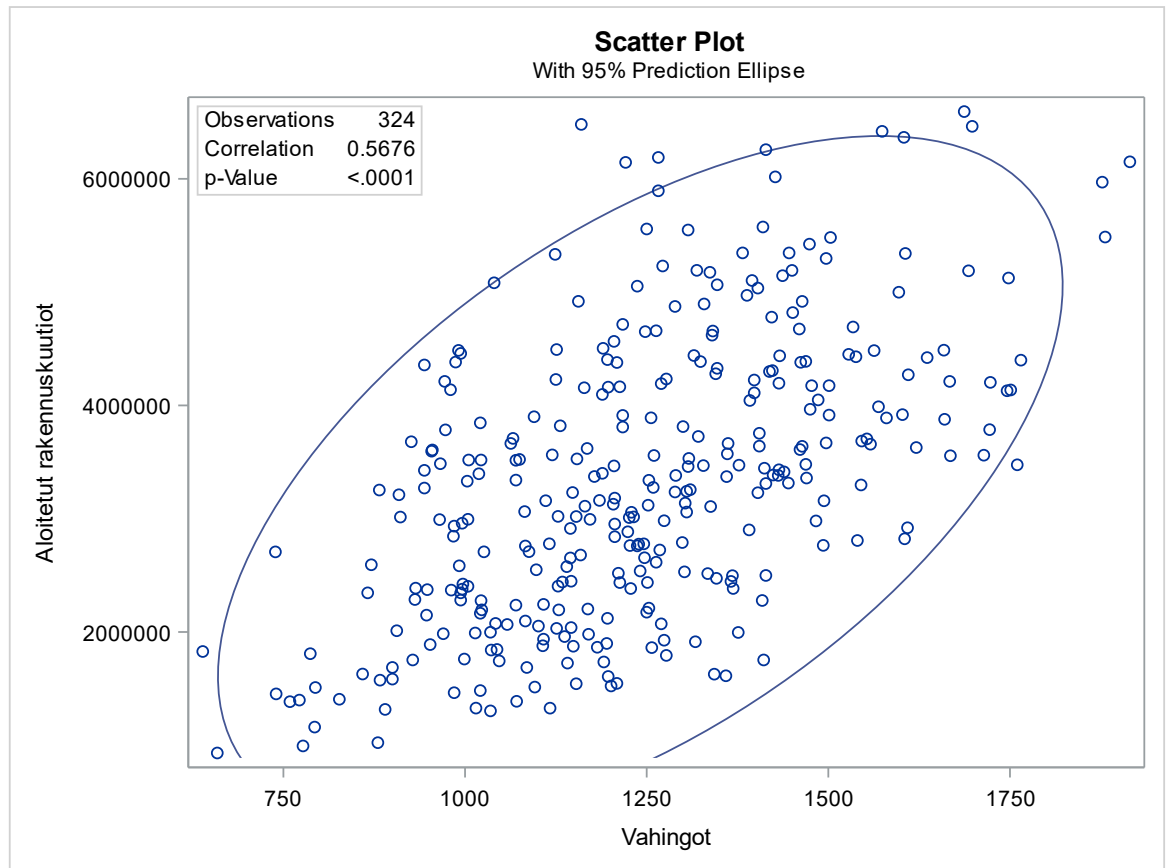
Toisena aikasarjana on Tilastokeskuksen aloitetut rakennushankkeet kuutioina (m³) kuukausittain vuosina 1995–2021. Käyttötarkoituksena on kaikki aloitetut rakennushankkeet ilman rajauksia ja rakennusvaihe on aloitetut rakennushankkeet. Aiemmin Tapaturmavakuutuskeskuksen käyttämässä ennustemallissa käytettiin liikevaihtoa selittävänä tekijänä, mutta nyt haluttiin kokeilla uuden ennusteen luomista uudella selittävällä tekijällä. Aloitetut miljoonat kuutiot -tieto on ladattu Tilastokeskuksen Statfinn-tietokannasta.

Työssä ei ole huomioitu vuoden 2005 alussa voimaan astunutta täyskustannusvastuun uudistamista. Tämän on arvioitu kasvattaneen tietyissä tapauksissa vahinkoilmoitusten määrää.

Taulukko 1. *Aloitettujen rakennuskuutioiden ja työtapaturmien välinen korrelaatio*

	Pearson Correlation Coefficients, N = 324 Prob > r under H0: Rho=0
Työpaikkatapaturmat - Aloitettut rakennuskuutiot	0.56761
Työpaikkatapaturmat - Aloitettut rakennuskuutiot	p < 0.0001

Tarkastellaan ensin yhdessä malleissa käytettyjä kahta aikasarjaa, eli aloitettuja rakennuskuutioita sekä sattuneita työtapaturmavahinkoja. Vahinkojen ja aloitettujen rakennuskuutioiden välillä voidaan todeta olevan yhteys, sillä taulukosta 1 tarkasteltuna Pearsonin korrelaatiokerroin saa arvon 0.5676 ja se on tilastollisesti merkitsevä (p-arvo < 0.05). Vahinkojen ja aloitettujen rakennuskuutioiden välillä näyttää olevan positiivinen yhteys. Aloitettujen kuutioiden lukumäärän kasvaessa myös vahinkojen lukumäärä kasvaa. Yhteyttä voi tarkastella myös visuaalisesti kuvasta 9. Mitä kapeampi kuvan ellipsi on, sitä vahvemmin muuttujat korreloivat.



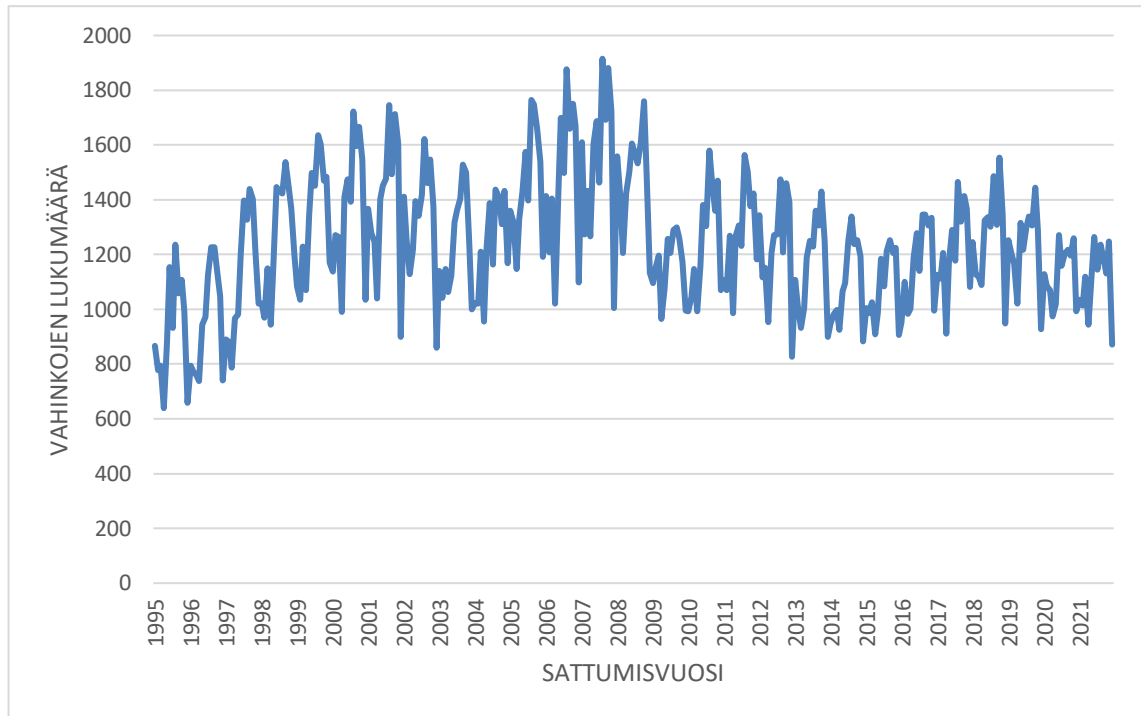
Kuva 9. Vahinkojen ja aloitettujen rakennuskuutioiden välinen korrelaatio

4.2.1 Työpaikkatapaturmat

Taulukko 2. Työpaikkatapaturmien tunnuslukuja

	Keskiarvo	Keskihajonta	Minimi	Maksimi	Lukumäärä
Työpaikkatapaturmat	1241	235.62	639	1915	324

Taulukossa 2 on esitetty työpaikkatapaturmien tunnuslukuja. Työpaikkatapaturmia sattuu rakennusteollisuudessa keskimäärin 1241 kappaletta kuukaudessa keskihajonnan ollessa 235,62. Pahimmillaan vahinkoja on sattunut 1915 kappaletta kuukaudessa, mutta parhaana kuukautena niitä on sattunut vain 639 kappaletta. Tutkimusaineisto koostuu 324 kuukaudesta.



Kuva 10. Työpaikkatapaturmien lukumäärät vuosina 1995–2021

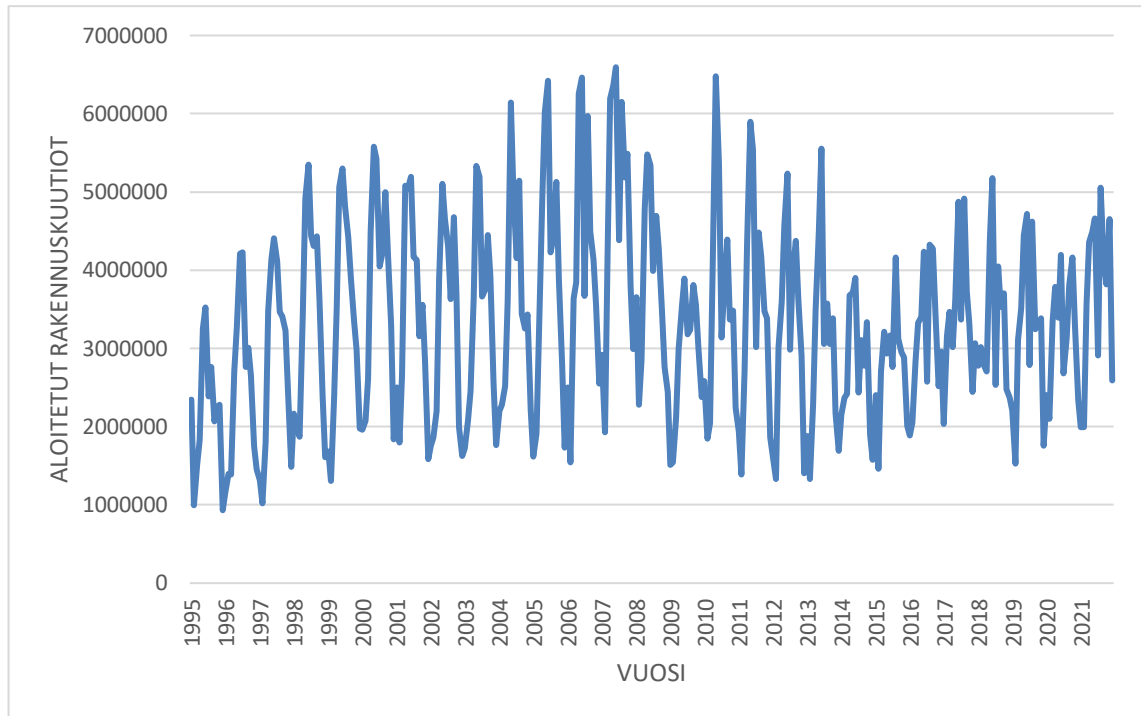
Työpaikkatapaturmien määrät vaihtelevat paljon kuukausitasolla (kuva 10) ja vaikuttaa siltä, että korkeimmat piikit osuvat kesäaikaan. Tämä on luontevaa alan kausiluonteisuudesta johtuen.

4.2.2 Aloitetut rakennushankkeet

Taulukko 3. Aloitetujen rakennushankkeiden kuutioiden tunnuslukuja

	Keskiarvo	Keskihajonta	Minimi	Maksimi	Lukumäärä
Aloitetut rakennuskuutiot	3334150	1233874.78	932049	6593809	324

Aloitetujen rakennushankkeiden kuutioissa on valtavaa vaihtelua (taulukko 3). Rakennushankkeita aloitetaan keskimäärin yli 3,3 miljoonan kuution edestä kuukausitasolla keskihajonnan ollessa yli 1,2 miljoonaa, eli kuukausitasolla aloitetuissa kuutioissa on suurta vaihtelua. Pienimmillään aloitettuja kuutiota on yli 932 000 kuutiota ja suurimmillaan lähes 6,6 miljoonaa aloitettua rakennushankkeen kuutiota.



Kuva 11. Aloitettujen rakennushankkeiden määrät aloitettuna kuutioina vuosina 1995–2021

Aloitettujen rakennushankkeiden määrät aloitettuna kuutioina vaihtelevat paljon kausitasolla (kuva 11).

4.3 Mallien soveltaminen aineistoon

Tutkimuksessa käytetyt LSTM- ja GRU-ennustemallit on rakennettu TensorFlow-ohjelmakirjaston avulla. TensorFlow on Googlen vuonna 2015 julkaisema syväoppimisen viitekehys. Tähän ohjelmakirjastoon on tullut uusia versioita ja TensorFlow on yksi laajemmin käytetyistä ohjelmakirjastoista.

Seuraavaksi tutustutaan sovellettujen mallien tarkempiin yksityiskohtiin, jotta malleja on helpompaa käyttää tulevaisuudessa ja mallin rakenteen muokkaus onnistuu tarvittaessa. Muutoksia on tarpeen tehdä, jos esimerkiksi aikasarjat pitenevät tai jos käytetään jotain muuta selittävää tekijää kuin rakennushankkeiden aloitettuja rakennuskuutioita.

Monet koneoppimisen algoritmit suoriutuvat paremmin, kun numeeriset selittävät muuttujat on skaalattu (Bishop, 1998). Tässä tutkimuksessa molemmat käytettävissä olevat muuttujat on normalisoitu arvovälille (0,1) MinMaxScaler-funktion avulla. Muuten neuroverkko voisi oppia mitä tahansa, sillä painotus on tässä tutkimuksessa käytettävässä aineistossa aloitettujen rakennushankkeiden kuutioiden, koska aloitettujen kuutioiden absoluuttinen lukumäärä on huomattavasti suurempi verrattuna sattuneisiin työtap-

turmiin millä tahansa tarkastelujaksolla. Yleisesti ottaen voidaan ajatella, että skaalaminen on hyvä käytäntö koneoppimisen algoritmien suoriutumisen kannalta. (Bishop, 1998)

Ennusteikkunan pituuden (Window Size) avulla määritellään, monellako edeltävällä arvolla testiaineistossa luodaan ennuste. Tämän jälkeen aineistoa käsitellään niin, että testiaineistolla tuotettu ennuste tulee käyttöön seuraavaan testiaineistoon. Tätä voidaan havainnollistaa käyttämällä ikkunan pituutena arvoa kuusi (6), työtapaturmavahinkojen lukumäärästä tietyllä ajanhetkellä lyhennettä t ja mallin tuottamasta ennusteesta tietyllä ajanhetkellä lyhennettä e :

$(t-5, t-4, t-3, t-2, t-1, t) \rightarrow e(1)$

$(t-4, t-3, t-2, t-1, t, e(1)) \rightarrow e(2)$

$(t-3, t-2, t-1, t, e(1), e(2)) \rightarrow e(3)$.

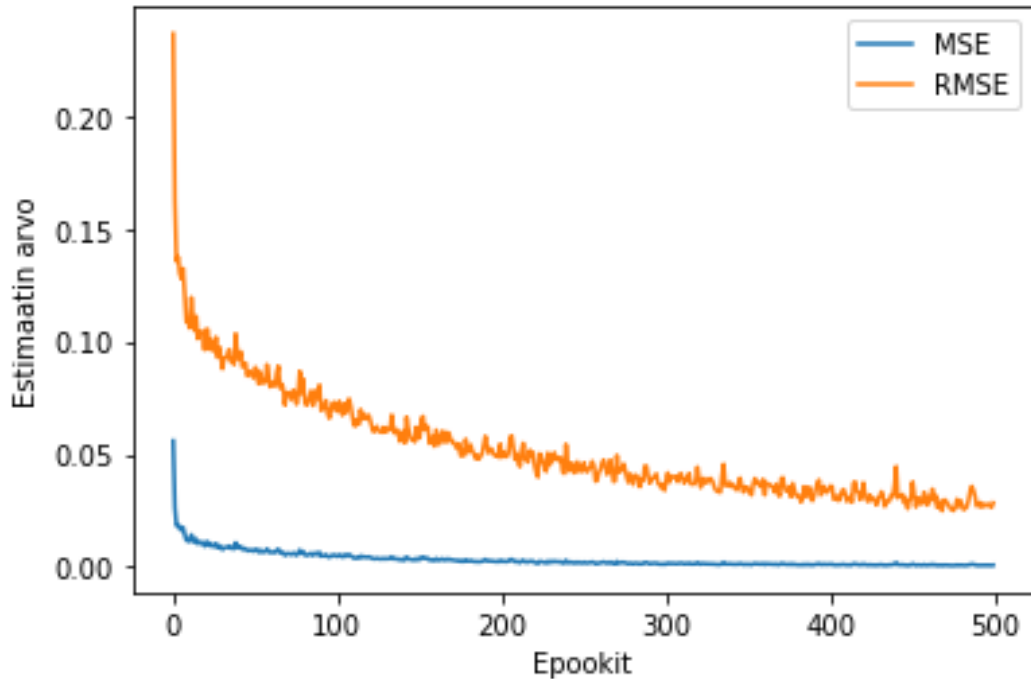
Jos ikkunan pituus saa esimerkiksi arvon kuusi (6), niin arvojoukko $(t-5, t-4, t-3, t-2, t-1, t)$ muodostaa ennusteen $e(1)$. Tämän jälkeen ennustearvo siirretään testidataan mukaan ja sitä hyödynnetään laskemaan seuraava ennustearvo $e(2)$. Näin tehtynä aina vanhin testiaineisto arvo tiputetaan pois testiaineistosta ja otetaan uusin ennuste aina mukaan. Tätä toimintatapaa sovelletaan sekä vahinkojen että aloitettujen rakennuskuutioiden ennustamiseen, koska vahinkojen tarkempi ennustaminen vaatii myös rakennuskuutioiden ennusteen hyödyntämistä. Rakennuskuutiot toimivat mallissa vain lisätietona, sillä tässä työssä tarkoituksena on vahinkojen ennustaminen.

Syväoppivissa neuroverkoissa käytetään erilaisia hyperparametreja verkon parantamiseksi. Yksi näistä on oppimisnopeus (Learning Rate). Syväoppivia neuroverkkoja opetetaan esimerkiksi stokastisen gradientin laskeutumisen (Stochastic Gradient Descent) algoritmilla. Se on optimointialgoritmi, joka päivittää neuroverkossa käytettäviä painoja. Se, miten paljon painoja päivitetään opetuksen aikana, kutsutaan oppimissuhteeksi. Oppimissuhde saa arvoja välillä (0.0, 1.0). Oppimissuhde määrittelee, kuinka nopeasti malli muuntautuu käsittelemään ongelmaa. Pienemmillä arvoilla tarvitaan käyttöön enemmän laskentakierroksia, koska muutoksia painoihin tehdään vähän kerrallaan jokaisella päivityskierroksella. Laskentakierrosta, jossa painokertoimien arvoja muutetaan, kutsutaan epookiksi (epoch). Suuremmilla oppimissuhteen arvoilla muutoksia tehdään enemmän jokaisella päivityskierroksella, mikä taas vaatii vähemmän epookkeja. Oppimissuhteen voidaan ajatella olevan yksi tärkeimmistä mallissa käytet-

tävistä hyperparametreista, joten on tärkeää valita se käytettävään malliin sopivaksi. (Brownlee, 2023)

Toisena hyperparametrina käytetään opetusdatan eräkokoja (Batch Size). Se määrittelee opetusdatajoukon, mikä esitellään mallille ja erän jälkeen mallin sisäisiä parametreja päivitetään.

Mallin rakentamista varten kerätty aineisto jaetaan kolmeen osa-aineistoon: opetusaineistoon (training dataset), testiaineistoon (testing dataset) sekä validointiaineistoon (validation dataset). Testaus- ja validointiaineistoilla varmistetaan mallin toimivuus aineistolla, jota ei ole hyödynnetty opetusvaiheessa. Koulutusvaiheen laskentaprosessin aluksi painokertoimet saattavat olla huonoja, mikä aiheuttaa epätarkkuutta ensimmäisillä laskentakierroksilla. Tämä johtuu siitä, että aluksi painokertoimien arvot ovat satunnaisesti valittuja. Neuroverkon painokertoimia voidaan päivittää vertaamalla neuroverkon ulostuloa valmiiksi tunnettuun eli todelliseen arvoon. Etukäteen valitulla kustannusfunktiolla voidaan laskea erotus neuroverkon ulostulon ja todellisen arvon välillä. Jos edellä mainitut arvot eroavat paljon, niin kustannusfunktion arvo on suuri. Neuroverkko toimii siinä vaiheessa hyvin, kun tunnettu arvo ja laskettu arvo ovat mahdollisimman lähellä toisiaan. Neuroverkon painokertoimien arvoja muutetaan pikkuhiljaa, jotta saadaan mahdollisimman hyvä lopputulos. Laskentakierroksella, jossa painokertoimien arvoja muutetaan, tarkoitetaan neuroverkon opettamista ja epookkeja voi laskennassa olla suuriakin määriä. Mallin tarkkuus siis paranee epookkien määrän kasvaessa. Toinen vertailun kohde on mallin virhe-kuvaaja (loss). Virhe-kuvaaja kuvastaa kustannusfunktion tulosta, joka pienenee laskennan edetessä. Jos lasketut arvot ja tunnetut arvot eroavat vain vähän, niin kustannusfunktion arvo on pieni. Näitä kuvaajia käytetään mallin toimivuuden tarkasteluun (kuva 12). (Kananen, 2019)



Kuva 12. Mallin tarkkuuden paraneminen

Yleisesti optimointialgoritmi käyttää edeltävien vaiheiden gradientteja ohjaamaan hakusuuntaa ja askelkokoa, jotta lähestytään nopeammin optimia (Dawani, 2020). Tensorflow-ohjelmakirjastossa optimointialgoritmeja on käytössä yli kymmenen erilaista. Mallissa käytetty Adam-optimointialgoritmi on laajalti tunnistettu yleisesti parhaaksi algoritmiksi.

Keskineliövirheen (Mean Squared Error - MSE) laskenta tapahtuu vertaamalla todellista arvoa ennusteen tuottamaan arvoon. Tässä tutkielmassa ennustavan mallin hyvyyden mittarina käytetään pienintä keskineliövirheen neliöjuuren arvoa (Root Mean Square Error - RMSE) (Kaava 1):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (1)$$

jossa \hat{y}_i tarkoittaa mallin tuottamaa ennustetta ja y_i todellista arvoa. (Gocheva-Ilieva, 2021)

Neuroverkkojen hyperparametrien muokkaaminen perustuu yleensä kokemukseen ja kokeilemiseen. Näin ollen optimaalisten arvojen löytäminen perustuu enemmän ja vähemmän testaamiseen (Kwon et al., 2020).

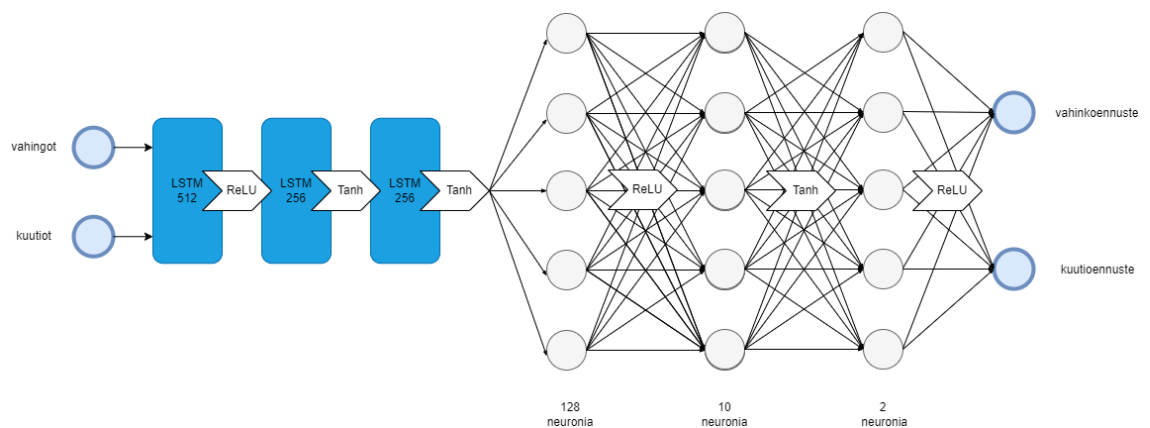
Neuroverkkoja joudutaan kouluttamaan useita kertoja sisäisten alkuarvojen satunnaisuudesta riippuen. Näistä malleista valitaan se, joka antaa pienimmän virheen opetus- ja validointidatan suhteen.

5. TULOKSET

Tässä tutkimuksessa käytettävät aikasarjat ovat niin lyhyeltä ajanjaksolta, että mallin rakentamisessa käytetään ainoastaan opetus- ja testiaineistoa. Tuloksissa esitellään käytetyt LSTM- ja GRU-mallit sekä niiden tuottamat ennusteet.

5.1 LSTM-malli

LSTM-malli (kuva 13) sisältää kolme neuroverkkokerrosta, joista ensimmäisessä oli 512 neuronia ja toisessa sekä kolmannessa kerroksessa 256 neuronia. Käytetyt aktivaatiofunktiot esitetään ohjelmointikoodissa liitteessä C. Malliin on lisätty kolme Dense-kerrosta LSTM-kerroksien jälkeen. Näissä kerroksissa jokainen neuroni on kytketty edeltävän kerroksen jokaiseen neuroniiin ja tällä tavoin pyritty heijastamaan ennusteaineistoa ja muodostamaan yhteyttä kerroksen ulostulon kanssa. Dense-kerroksissa on käytössä eri määrä neuroneja ja eri aktivaatiofunktioita. Ensimmäisessä kerroksessa käytetään 128 neuronia ja aktivaatiofunktiona käytetään ReLU-funktiota. Toisessa kerroksessa on 10 neuronia ja Tanh-aktivaatiofunktio kun taas kolmannessa kerroksessa on käytössä 2 neuronia ja ReLU-aktivaatiofunktio.

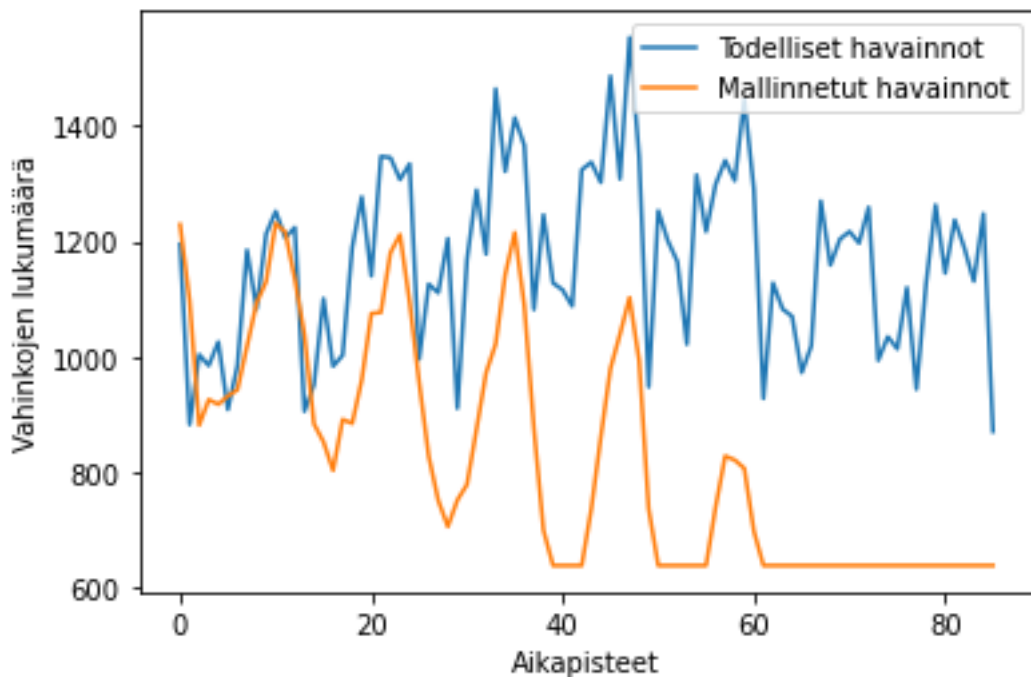


Kuva 13. LSTM-neuroverkon rakenne

Käytetyssä mallissa LSTM-kerrosten väliin on lisätty pudotuskerros (Dropout layer). Tämä on eräs suosituimmista säännöllistämistekniikoista, joita käytetään mallin ylisovittamisen estämiseksi. Verkolla on kuitenkin taipumus oppia sellaisia datan ominaisuuksia, joita ei voida yleistää ja se voi johtaa mallin ylisovittamiseen. Pudottaminen palaut-

taa osan piilotetun yksikön ulostuloista sattumanvaraisesti nolaksi perustuen annettuun arvoon vähentääkseen neuroverkon ylioppimista. Tässä tutkimuksessa pudotuskerros on yhdistetty kunkin LSTM-kerroksen väliin oppimisvaiheen helpottamiseksi ja nopeuttamiseksi. (Mahjoub et al., 2022)

Kuvassa 14 esitetään, miten huonosti malli voi toimia opetuskierrosten jälkeen. Kuvassa havainnollistetaan, miten malli käyttäytyy opetuskierrosten jälkeen verrattaessa todelliseen aineistoon. Vahinkojen todelliset arvot vaihtelevat noin 900 ja 1550 vahingon välissä, mutta malli ennustaa noin 20 havainnon jälkeen vahinkojen lukumääräksi huomattavasti pienempiä arvoja, mikä ei tietysti kuvasta hyvin todellista tilannetta. Mallin kokeiluvaiheessa, kun parametrien määrittäminen ja verkon arkkitehtuuri ovat vielä kehitysvaiheessa, niin mallinnetut havainnot eivät vastaa todellisia havaintoja.

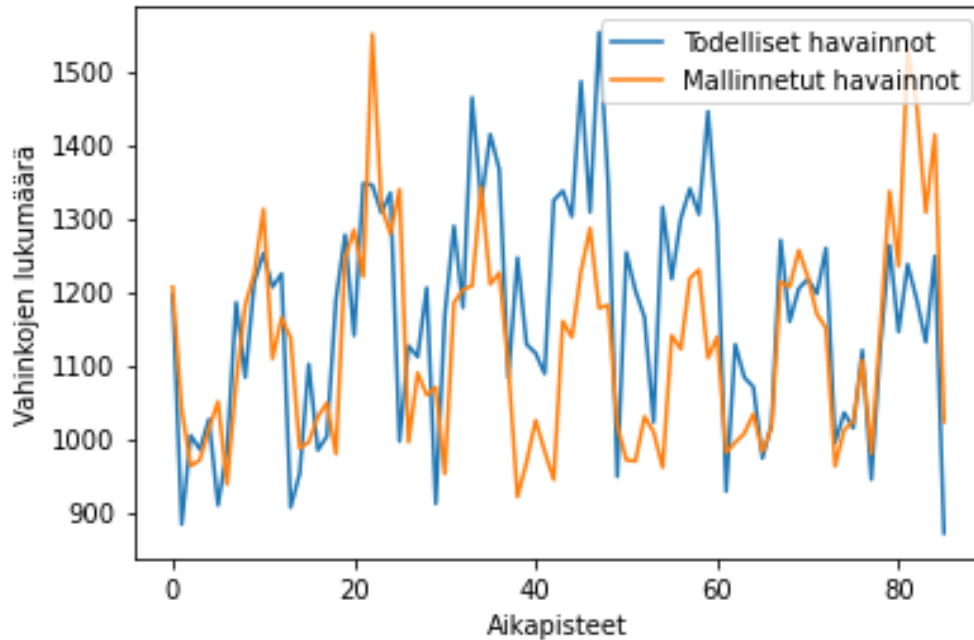


Kuva 14. Huonosti oppiva malli

Huonosti oppiva malli ei auta ennustamisessa, joten testaamista on jatkettava eri arkkitehtuurilla ja parametreilla. Kuvan 15 hyvin oppivassa mallissa käytetään seuraavia hyperparametreja:

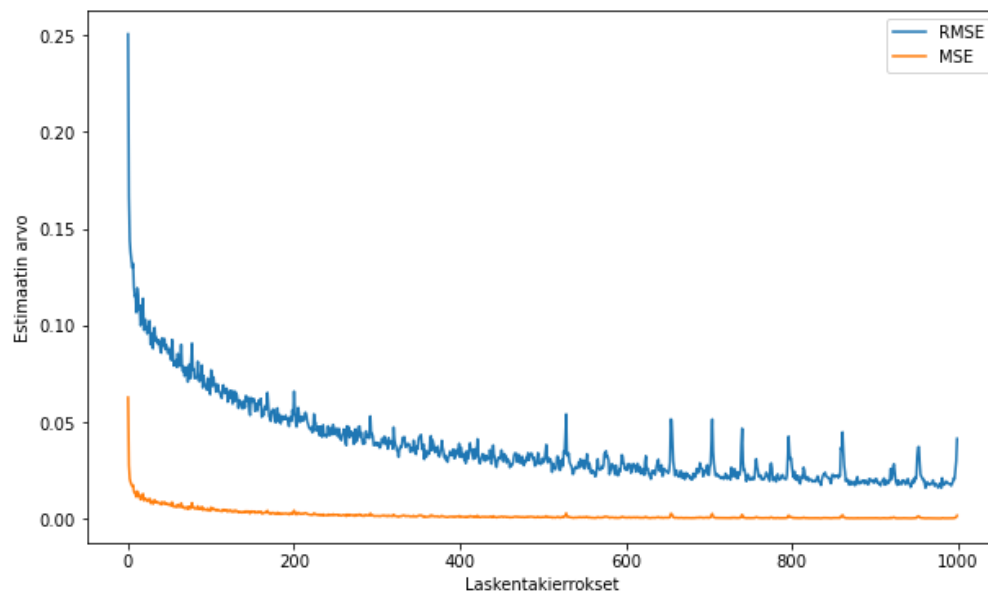
- learning_rate = 0.001
- batch_size = 4
- epochs = 1000.

Mallissa opetusvaihe on jo paljon parempi kuin edellä mainitussa huonossa esimerkissä. Parannellun mallin tuottama kuvaaja seuraa opetusaineistona käytettyjä todellisia arvoja huomattavasti paremmin (kuva 15). Pitkälläkin tilastoaineistolla opetetun mallin voi olla vaikeaa seurata erittäin nopeita muutoksia



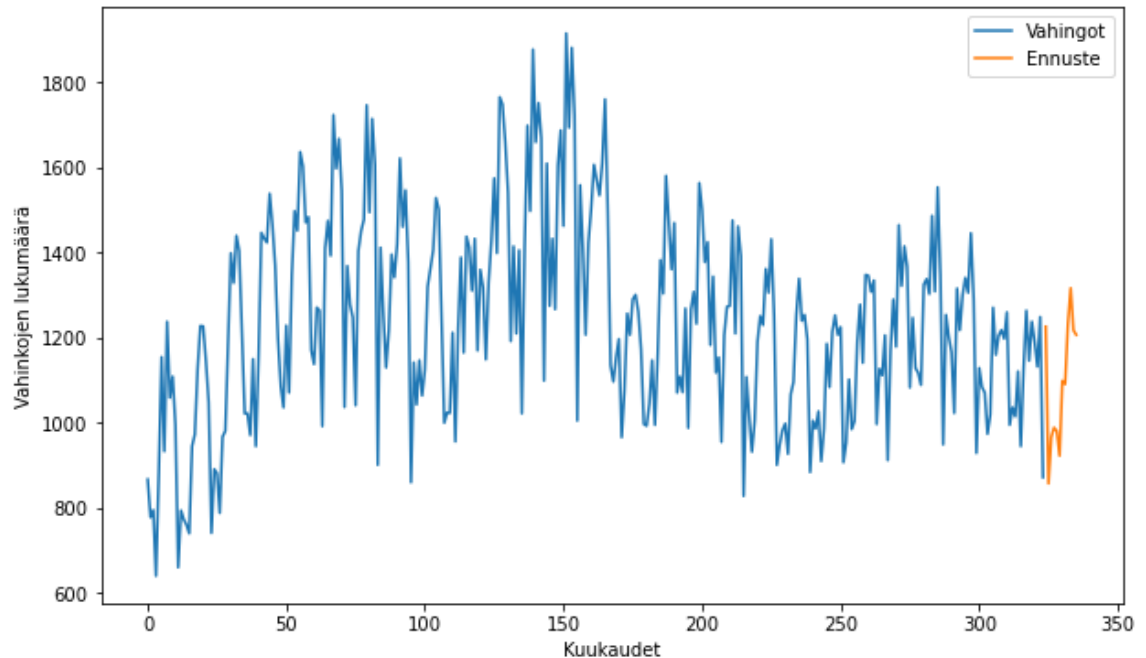
Kuva 15. LSTM-mallin tuottama mallinnus vs. opetusaineisto

LSTM-mallin perusteella lasketut keskineliövirheet sekä juuri-keskineliövirheet paranevat laskentakierrosten edetessä (kuva 16).



Kuva 16. LSTM-mallin keskineliövirheen ja juuri-keskineliövirheen kehittyminen

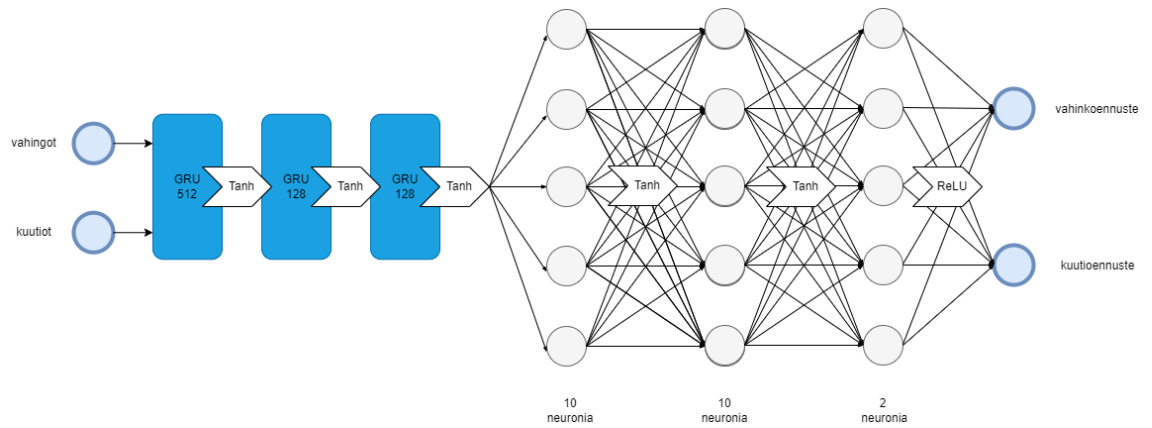
LSTM-malli tuottaa järkevän näköisen ennusteen (kuva 17). Malli tuottaa 12 kuukauden päähän ennusteen tulevaisuudessa sattuvista työtapaturmista rakennusteollisuudessa. Yhteenlaskettu 12 kuukauden työtapaturmien määrä on 13095.



Kuva 17. LSTM-mallin tuottama ennuste

5.2 GRU-malli

GRU-neuroverkossa käytetään aktivaatiofunktiona kahta eri aktivaatiofunktioita. Käytetyt aktivaatiofunktiot esitetään ohjelmointikoodissa liitteessä D. Ensimmäisessä piilokerroksessa on 512 neuronia ja aktivaatiofunktiona käytetään Tanh-funktiota. Tämän jälkeen lisätään pudotuskerros. Pudotuskerros deaktivoi osan neuroneista arvolla 0.2. Toisessa piilokerroksessa on 128 neuronia ja aktivaatiofunktiona käytetään Tanh-funktiota. Kolmannessa piilokerroksessa on 128 neuronia ja käytössä on Tanh-aktivaatiofunktio. Näiden kummankin kerroksen jälkeen on jo edellä mainittu pudotuskerros. Lisäksi malliin on lisätty perään kolme Dense-kerrosta. Dense-kerroksissa on käytössä eri määrä neuroneja ja eri aktivaatiofunktioita. Ensimmäisessä kerroksessa ja toisessa kerroksessa käytetään 10 neuronia ja aktivaatiofunktiona käytetään Tanh-funktiota, kun taas kolmannessa kerroksessa on käytössä 2 neuronia ja ReLU-aktivaatiofunktio. Yksinkertaistetun mallin rakenne käy ilmi kuvasta 18.

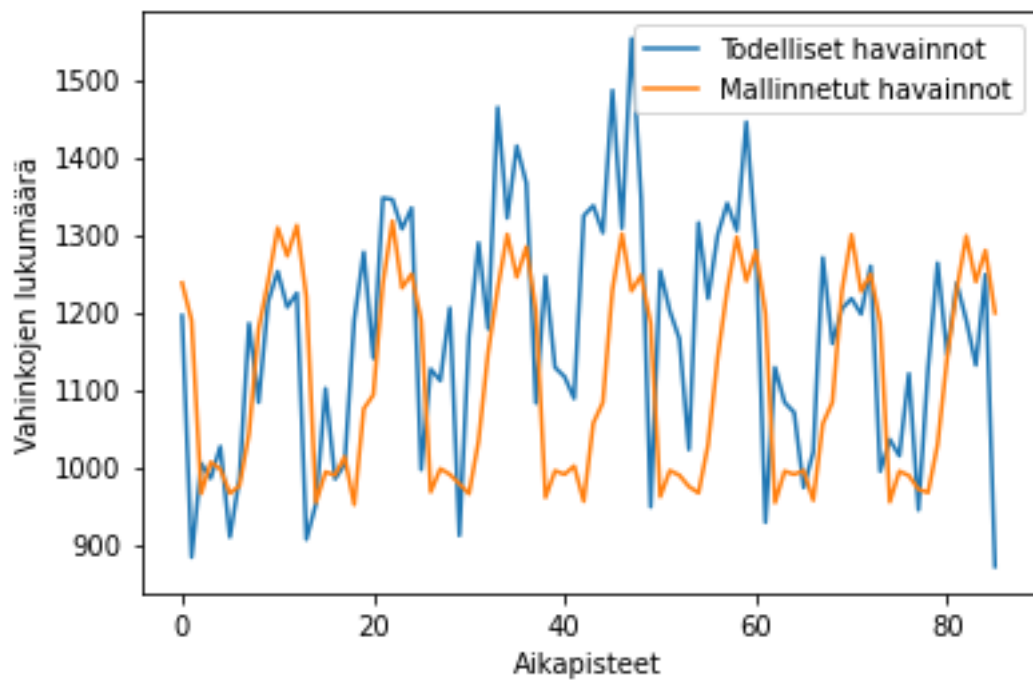


Kuva 18. GRU-neuroverkon rakenne

Mallissa on käytössä seuraavat hyperparametrit:

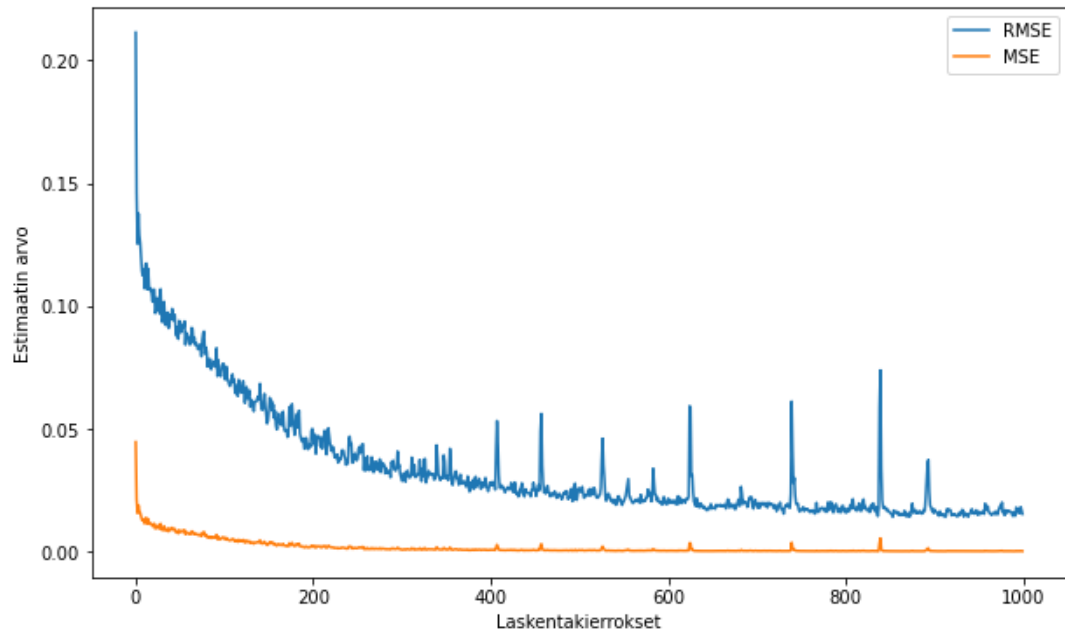
- learning_rate = 0.001
- batch_size = 4
- epochs = 200.

GRU-mallissa opetusvaihe on jälleen paljon parempi mitä edellä mainitussa huonossa esimerkissä (kuva 14). Mallin tuottama kuvaaja mukaillee opetusaineistona käytettyjä todellisia arvoja huomattavasti paremmin (kuva 19).



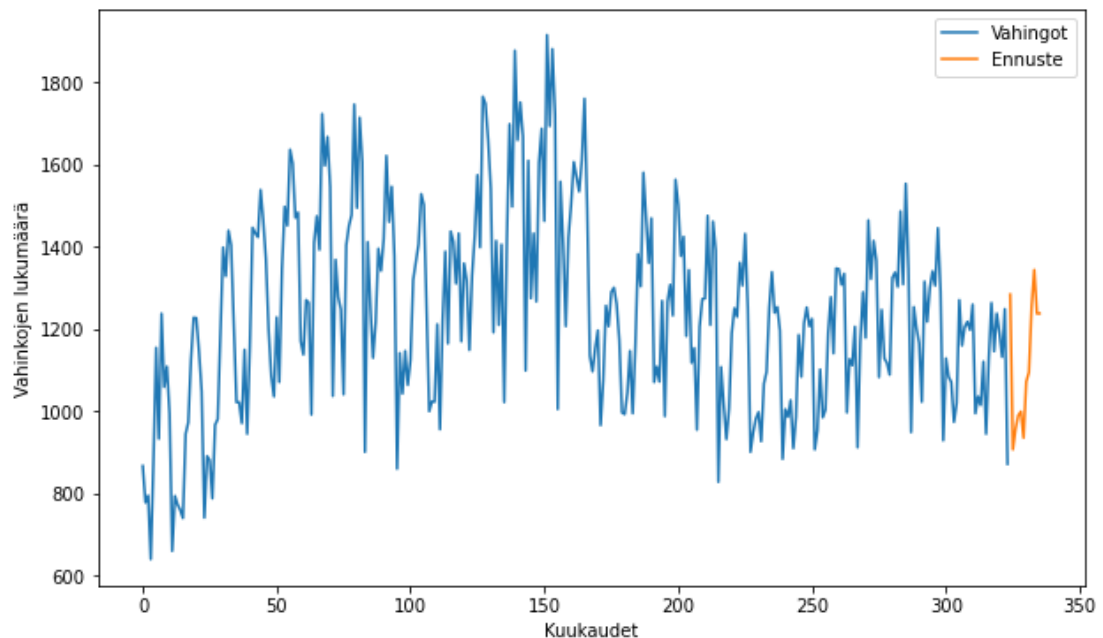
Kuva 19. GRU-mallin tuottama mallinnus vs. opetusaineisto

GRU-mallin perusteella lasketut keskineliövirheet sekä juuri-keskineliövirheet paranevat laskentakierrosten edetessä (kuva 20).



Kuva 20. GRU-mallin keskineliövirheen ja juuri-keskineliövirheen kehittyminen

GRU-malli tuottaa järkevän näköisen ennusteen (kuva 21). Malli tuottaa 12 kuukauden päähän ennusteen tulevaisuudessa sattuvista työtapaturmista rakennusteollisuudessa. Yhteenlaskettu 12 kuukauden työtapaturmien määrä on 13304.



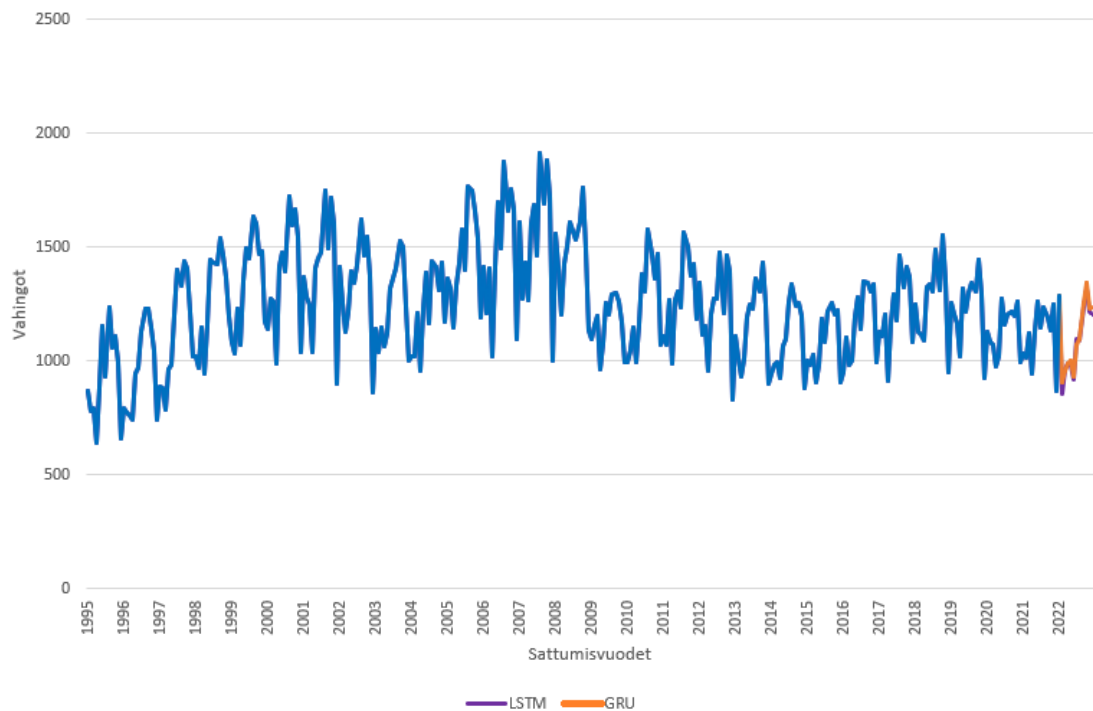
Kuva 21. GRU-mallin tuottama ennuste

5.3 Mallien vertailu

Taulukossa 4 on esitetty juuri-keskineliövirheen (RMSE) arvot molemmille malleille kuvista 15 ja 19 sekä yhteenlaskettu ennuste 12 kuukauden päähän tulevaisuudessa tapahtuville työtapaturmille rakennusteollisuudessa. Tulosten perusteella LSTM-malli suoriutuu hieman paremmin kuin GRU-malli, kun verrataan juuri-keskineliövirhettä. Ennusteen arvo poikkeaa jonkin verran mallien välillä (LSTM 13095 vs. GRU 13304). Molempien mallien tuottamat ennusteet vaikuttavat hyvin samanlaisilta (kuva 22): ennusteen mukaan vahinkojen määrä nousee ennustejakson loppua kohden ja käyrän muoto on hyvin samankaltainen.

Taulukko 4. LSTM- ja GRU-mallin tulosten vertaaminen

	Juuri-keskineliövirhe (RMSE)	Ennuste (12kk)
LSTM	153	13095
GRU	160	13304



Kuva 22. LSTM- ja GRU-mallien tuottamien ennusteiden vertaaminen

6. JOHTOPÄÄTÖKSET

Tutkimuksen tarkoituksena oli tuottaa vuoden päähän ennuste rakennusteollisuudessa tapahtuvista työtatapaturmavahingoista. Aineistossa oli mukana kaikki palkansaajien työpaikkatapaturmat rakentamisen päätoimialan talonrakentaminen sekä erikoistunut rakennustoiminta alaluokissa sattumisvuosina 1995–2021. Aineistossa ei ollut mukana työmatkalla sattuneita työtapaturmia eikä yrittäjille sattuneita työtapaturmia. Organisaatiossa aiemmin käytetty VAR-malli ennustaa kaikkia työtapaturmia, joten sitä ei ole mielekästä verrata tässä tutkielmassa rakennettuihin neuroverkkomalleihin.

Tutkimuksessa tuotettiin 12 kuukauden ennuste LSTM- ja GRU-neuroverkkomallien avulla. Mallit erosivat tuloksiltaan hieman, sillä LSTM-mallin juuri-keskineliövirhe oli 153 kun taas GRU-mallin 160. LSTM-malli ennustaa 13095 työtapaturmaa ja GRU-malli 13304 työtapaturmaa rakennusteollisuuden alalle vuonna 2022. Viimeisen kymmenen vuoden aikana on sattunut keskimäärin 12 859 vahinkoa ja viimeisen viiden vuoden aikana on sattunut keskimäärin 12 973 vahinkoa. Voidaan siis ajatella, että tuotetut ennusteet ovat linjassa toteutuneiden vahinkomäärien kanssa. Tuorein sattumisvuosi, eli sattumisvuosi 2022 valmistuu kesällä 2023, joten tulevien kuukausien aikana on mahdollista verrata, kuinka hyvin käytetyt mallit osasivat ennustaa tulevia vahinkoja. LSTM-malli ennusti hieman lähemmäs menneiden vuosien keskiarvoja. Tämä johtuu luultavasti siitä, että malli on monimutkaisempi ja tuottaa näin ollen paremman ennusteen.

Tässä tutkimuksessa käytetyt mallit eivät välttämättä ole parhaat löydettävissä olevat mallit, vaan malleihin valitut parametrien arvot on pyritty löytämään testaamalla laajasti erilaisilla arvoilla. Parhaiden parametrien löytäminen perustuu kokemukseen ja testaamiseen. Tutkielmaa varten luodut mallit ovat käyttökelpoisia myös muissa sovelluskoh-teissa, kunhan mallin parametrit säädetään ja malli opetetaan uudestaan. Ovatko käytetyt LSTM- ja GRU-mallit parhaat mahdolliset ylipäätään tähän? Voidaanko käyttää kaksisuuntaisia malleja? Malleja voidaan hienosäätää joko Keras-ohjelmakirjaston tai tässä työssä käytetyn Tensorflow-ohjelmakirjaston avulla. Ohjelmakirjastoilla on erilaisia parametreja, joten toisella kirjastolla saatettaisiin päästä parempiin tuloksiin. Tiivistäen, neuroverkoissa on vielä paljon opeteltavaa ja näin ollen malleissa myös parannettavaa. Tässä tutkimuksessa on vasta hieman perehdytty aiheeseen.

Aineistoa voidaan myös tarkastella kriittisesti. Kummassakin tutkimuksessa käytetty aikasarja voi erota vuositasolla, sillä asioita voi mitata monella tavalla. Voidaanko varmistua, että läpi koko ajanjakson tiedot on toimitettu samalla tavalla ja yhteneväisesti? Ilmoitetuissa aloitettujen rakennushankkeiden kuutioissa voi olla eroja, eli onko kaikki aloitetut kuutiot varmasti toimitettu tilastoon. Kaikki edellä olevat asiat vaikuttavat siihen, onko aineisto vertailukelpoista eri vuosien välillä.

Suurimman suhdannevaihtelun omaavat päätoimialat (rakentaminen, teollisuus, kuljetus ja kauppa) voitaisiin jatkossa ennustaa näillä menetelmillä ja summata sitten ennusteet yhteen niiden toimialojen kanssa, joissa vahinkojen vaihtelu on vähäisempää. Tätä arviota voidaan verrata työtuntien ennusteella ja asiantuntijalaitosten muodostamalla suhdanne-ennusteella muodostettuun asiantuntija-arvioon. Työssä esitettyjä menetelmiä hyödyntäen on tarkoitus tehdä ennuste syksyllä 2023.

Loppuvuodesta 2022 rakennusteollisuus on kääntynyt taantumaan korkotason nousun ja elvytyksen loppumisen myötä. Nämä asiat aiheuttavat todennäköisesti viive-eroa. Kaikkia aloitettuja rakennuskuutioita ei välttämättä rakenneta niin nopeasti valmiiksi. Näin työn intensiteetti ja siinä sattuvat tapaturmat eivät esiinny vakioviiveellä rakennusprojektin aloittamisesta.

Mallissa käytetty ajanjakso on tavallaan lyhyt, mutta työelämän kehityksen kautta tarkasteltuna ajanjakso onkin erittäin pitkä. Työ muuttunut paljon, sillä raskaita töitä tehdään paljon enemmän nykyään koneilla. Erilaiset selkä-, olkapää- ja kyynärpäävammat ovat selkeästi vähentyneet.

Jatkotutkimusaiheena voisi suorittaa tutkimusasetelman kääntämistä toisinpäin. Olisi mielenkiintoista selvittää, miten aloitetut rakennuskuutiot seuraavat sattuneita työtapaturmia. Tässä tarvitsisi luultavasti käyttää viivettä muuttujien välillä. Olisi mielekästä testata LSTM- ja GRU-malleja uudella aineistolla, jos tulevaisuudessa löytyy jotain aineistoa, mikä sopisi paremmin aloitettujen rakennushankkeiden kuutioiden sijasta mallissa käytettäväksi.

Vuonna 2017 julkaistussa artikkelissa esiteltiin niin sanottu enkoodaus-dekoodaus -arkkitehtuuri (encoder-decoder), joka perustuu huomiointikerrokseen. Kirjoittajat kutsuvat tätä nimellä "Transformer". Tätä niin sanottua Transformer-neuroverkkoarkkitehtuuria voisi harkita tulevaisuudessa käytettäväksi työtapaturmien ennustamiseen. (Utkarsh, 2022)

7. LÄHTEET

- Bhagwat, K., & Delhi, V. S. K. (2021). Review of construction safety performance measurement methods and practices: a science mapping approach [Article]. *International Journal of Construction Management*, *ahead-of-print*(ahead-of-print), 1–15. <https://doi.org/10.1080/15623599.2021.1924456>
- Bhagwat, K., Delhi, V. S. K., & Nanthagopalan, P. (2022). Construction safety performance measurement using a leading indicator-based jobsite safety inspection method: case study of a building construction project [Article]. *International Journal of Occupational Safety and Ergonomics*, *ahead-of-print*(ahead-of-print), 1–12. <https://doi.org/10.1080/10803548.2021.2012350>
- Bishop, C. M. (1998). *Neural networks for pattern recognition* (Repr.) [Book]. Clarendon.
- Brownlee, J. (2023, May 11). *Understand the Impact of Learning Rate on Neural Network Performance*. <https://Machinelearningmastery.Com/Understand-the-Dynamics-of-Learning-Rate-on-Deep-Learning-Neural-Networks/>. <https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/>
- Cho, K., Bart van Merriënboer, Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation [Document]. *ArXiv.Org*.
- Dawani, J. (2020). *Hands-On Mathematics for Deep Learning* (1st edition) [Book]. Packt Publishing.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM [Article]. *Neural Computation*, *12*(10), 2451–2471. <https://doi.org/10.1162/089976600300015015>
- Gocheva-Ilieva, S. (2021). *Statistical Data Modeling and Machine Learning with Applications* (S. Gocheva-Ilieva, Ed.) [Book]. MDPI - Multidisciplinary Digital Publishing Institute.
- Graves, A., Liwicki, M., Fernandez, S., Bertolami, R., Bunke, H., & Schmidhuber, J. (2009). A Novel Connectionist System for Unconstrained Handwriting Recognition [Article]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(5), 855–868. <https://doi.org/10.1109/TPAMI.2008.137>
- Hallowell, M. R., Bhandari, S., & Alruqi, W. (2020). Methods of safety prediction: analysis and integration of risk assessment, leading indicators, precursor analysis, and safety climate [Article]. *Construction Management and Economics*, *38*(4), 308–321. <https://doi.org/10.1080/01446193.2019.1598566>
- Kananen, H. (2019). *Tekoäly: bisneksen uudet työkalut* (H. Puolitaival, Ed.) [Book]. Alma Talent Oy.
- Kwon, B.-S., Park, R.-J., & Song, K.-B. (2020). Short-Term Load Forecasting Based on Deep Neural Networks Using LSTM Layer [Article]. *Journal of Electrical Engineering & Technology*, *15*(4), 1501–1509. <https://doi.org/10.1007/s42835-020-00424-7>
- Liu, W., Pan, J., Ren, Y., Wu, Z., & Wang, J. (2020). Coupling prediction model for long-term displacements of arch dams based on long short-term memory network [Article]. *Structural Control and Health Monitoring*, *27*(7), n/a. <https://doi.org/10.1002/stc.2548>
- Liu, X., Lin, Z., & Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA - A comparison against GRU and LSTM [Article]. *Energy (Oxford)*, *227*, 120492. <https://doi.org/10.1016/j.energy.2021.120492>
- Long, L. (2022). *Beginning deep learning with TensorFlow: work with Keras, MNIST data sets, and advanced neural networks* (X. Zeng, Ed.) [Book]. Apress L. P. <https://doi.org/10.1007/978-1-4842-7915-1>
- Loy, J. (2019). *Neural network projects with Python: the ultimate guide to using Python to explore the true power of neural networks through six projects* (1st edition) [Book]. Packt Publishing Ltd.














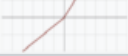






- Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., & Delahoche, L. (2022). Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks [Article]. *Sensors (Basel, Switzerland)*, 22(11), 4062. <https://doi.org/10.3390/s22114062>
- Pecican, E. S. (2010). Forecasting based on open VAR model [Article]. *Romanian Journal of Economic Forecasting*, 13(1), 59–69.
- Rakennusteollisuus ry A. (2023). *Tilastot ja suhdanteet*. <https://www.Rt.Fi/Tietoa-Alasta/Talous-Tilastot-Ja-Suhdanteet/>. <https://www.rtfi.fi/Tietoa-alasta/Talous-tilastot-ja-suhdanteet/>
- Rakennusteollisuus ry B. (2023). *Työturvallisuus*. <https://www.Rt.Fi/Tietoa-Alasta/Tyoturvallisuus/>. <https://www.rtfi.fi/Tietoa-alasta/Tyoturvallisuus/>
- Rantala, Jukka. (2009). *Vakuutusoppi* (11. uud. p.) [Book]. Finva.
- Salo, K. (2015). *Työtäpaturma ja ammattitauti* [Book]. Finva.
- Saloheimo, J. (2016). *Työturvallisuus: perusteet, vastuu ja oikeusturva* (3., uudistettu pa...) [Book]. Talentum Pro.
- Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM [Article]. *Chaos, Solitons and Fractals*, 140, 110212–110212. <https://doi.org/10.1016/j.chaos.2020.110212>
- Suomen Pankki. (1989). *Vektoriautoregressiiviset mallit*. https://helda.helsinki.fi/bof/bitstream/handle/123456789/15828/D_69.pdf?sequence=1&isAllowed=y
- Tapaturmavakuutuskeskus. (2020). *Työtäpaturmavakuutus numeroina 2020*. <https://api.tyotapaturmatieto.fi/file-store/0-292880-553528>
- Tapaturmavakuutuskeskus. (2022a, January). *Rakentamisen toimialalla pitkään jatkunut työpaikatäpaturmien taajuuden laskutrendi taittuu, taajuus kääntyi jo nousuun 2021*. <https://api.tyotapaturmatieto.fi/file-store/0-308359-581116>
- Tapaturmavakuutuskeskus. (2022b). *TOT-tutkinnan käsikirja*. <https://www.tvk.fi/tyoturvallisuus/tyopaikkaonnettomuuksien-tutkinta/tot-tutkinta/>
- Tilastokeskus. (2022). *Tutkimuksessa käytetty aineisto*. https://pxdata.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__ras/statfin_ras_pxt_12fy.px/
- Työterveyslaitos. (2019, October 2). *Rakennusalan työturvallisuuden kehitys*. NOLLA TAPATURMAA RAKENNUSTEOLLISUUDESSA 2020 -HANKE. https://www.rakennusteollisuus.fi/globalassets/tyoturvallisuus/2020_sekalainen/raportti_fin_al.pdf?utm_source=eJokka+2/2020+%7c+Aiheina+mm.+nolla+tapaturmaa,+turvallisuusvii kko,+koronaohjeita,+hiljainen+hetki&utm_medium=email&utm_campaign=eJokka
- TyTAL. (2015). *Tapaturma- ja työturvallisuuslaki (TyTAL)*. <https://www.Finlex.Fi/Fi/Laki/Ajantasa/2015/20150459>. <https://www.finlex.fi/fi/laki/ajantasa/2015/20150459>
- Utkarsh, A. (2022). *Transformer Neural Networks: A Step-by-Step Breakdown*. <https://Builtin.Com/Artificial-Intelligence/Transformer-Neural-Network>. <https://builtin.com/artificial-intelligence/transformer-neural-network>
- Zarzycki, K., & Ławryńczuk, M. (2021). LSTM and GRU Neural Networks as Models of Dynamical Processes Used in Predictive Control: A Comparison of Models Developed for Two Chemical Reactors [Article]. *Sensors (Basel, Switzerland)*, 21(16), 5625. <https://doi.org/10.3390/s21165625>
- Zazo, R., Lozano-Diez, A., Gonzalez-Dominguez, J., Toledano, D. T., & Gonzalez-Rodriguez, J. (2016). Language Identification in Short Utterances Using Long Short-Term Memory (LSTM) Recurrent Neural Networks [Article]. *PloS One*, 11(1), e0146917–e0146917. <https://doi.org/10.1371/journal.pone.0146917>

LIITE A: TYÖPAIKKATAPATURMAT SATTUMISVUOSITTAIN

Päätoimiala	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
C Teollisuus	21139	21596	19879	17994	16930	15651	15313	16119	15989	15626	13449
F Rakentaminen	14861	15188	14614	13908	13112	12879	13797	14573	15101	14711	13489
G Tukku- ja vähittäiskauppa	11451	11889	11769	11298	10914	10547	10759	10561	10691	10823	9433
H Kuljetus ja varastointi	9213	9450	9129	8555	7956	7702	7826	7759	7617	7694	5930
I Hallinto- ja tukipalvelutoiminta	7273	8074	8038	7414	7655	7864	8271	9446	9745	10282	8026
Q Terveys- ja sosiaalipalvelut	3500	3760	3970	4164	4245	4408	4715	4865	5294	5855	5651
Z Kuntasektori	19901	20650	20273	20800	20069	20565	21065	21631	20883	20665	17212
Muut	17084	17647	17732	17407	17195	16780	16871	17272	16954	17475	13416
Yhteensä	104422	108254	105404	101540	98076	96396	98617	102226	102274	103131	86606

Palkansaajien työpaikkatapaturmat päätoimialan mukaan sattumisvuosittain (TVK)

LIITE B: LISTAUS AKTIVAATIOFUNKTIOISTA

Name	Plot	Equation	Derivative (with respect to x)
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a. Sigmoid or Soft step)		$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$ ^[1]	$f'(x) = f(x)(1 - f(x))$
TanH		$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2 + 1}$
ElliotSig ^{[9][10][11]} Softsign ^{[12][13]}		$f(x) = \frac{x}{1 + x }$	$f'(x) = \frac{1}{(1 + x)^2}$
Inverse square root unit (ISRU) ^[14]		$f(x) = \frac{x}{\sqrt{1 + \alpha x^2}}$	$f'(x) = \left(\frac{1}{\sqrt{1 + \alpha x^2}} \right)^3$
Inverse square root linear unit (ISRLU) ^[14]		$f(x) = \begin{cases} \frac{x}{\sqrt{1 + \alpha x^2}} & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \left(\frac{1}{\sqrt{1 + \alpha x^2}} \right)^3 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Square Nonlinearity (SQNL) ^[11]		$f(x) = \begin{cases} 1 & : x > 2.0 \\ x - \frac{x^2}{4} & : 0 \leq x \leq 2.0 \\ x + \frac{x^2}{4} & : -2.0 \leq x < 0 \\ -1 & : x < -2.0 \end{cases}$	$f'(x) = 1 \mp \frac{x}{2}$
Rectified linear unit (ReLU) ^[15]		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Leaky rectified linear unit (Leaky ReLU) ^[16]		$f(x) = \begin{cases} 0.01x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0.01 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parametric rectified linear unit (PReLU) ^[17]		$f(\alpha, x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(\alpha, x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Randomized leaky rectified linear unit (RReLU) ^[18]		$f(\alpha, x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$ ^[3]	$f'(\alpha, x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus ^[23]		$f(x) = \ln(1 + e^x)$	$f'(x) = \frac{1}{1 + e^{-x}}$
Bent identity		$f(x) = \frac{\sqrt{x^2 + 1} - 1}{2} + x$	$f'(x) = \frac{x}{2\sqrt{x^2 + 1}} + 1$
SoftExponential ^[26]		$f(\alpha, x) = \begin{cases} -\frac{\ln(1 - \alpha(x + \alpha))}{\alpha} & \text{for } \alpha < 0 \\ x & \text{for } \alpha = 0 \\ \frac{e^{\alpha x} - 1}{\alpha} + \alpha & \text{for } \alpha > 0 \end{cases}$	$f'(\alpha, x) = \begin{cases} \frac{1}{1 - \alpha(x + \alpha)} & \text{for } \alpha < 0 \\ e^{\alpha x} & \text{for } \alpha \geq 0 \end{cases}$
Soft Clipping ^[27]		$f(\alpha, x) = \frac{1}{\alpha} \log \frac{1 + e^{\alpha x}}{1 + e^{\alpha(x-1)}}$	$f'(\alpha, x) = \frac{1}{2} \sinh\left(\frac{p}{2}\right) \operatorname{sech}\left(\frac{px}{2}\right) \operatorname{sech}\left(\frac{p}{2}(1-x)\right)$
Sinusoid ^[28]		$f(x) = \sin(x)$	$f'(x) = \cos(x)$
Sinc		$f(x) = \begin{cases} 1 & \text{for } x = 0 \\ \frac{\sin(x)}{x} & \text{for } x \neq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x = 0 \\ \frac{\cos(x)}{x} - \frac{\sin(x)}{x^2} & \text{for } x \neq 0 \end{cases}$
Gaussian		$f(x) = e^{-x^2}$	$f'(x) = -2xe^{-x^2}$

Lähde: Wikipedia

LIITE C: LSTM-NEUROVERKKO

LSTM-neuroverkko rakennusteollisuuden työtaturmien ennustamiseen. Aineiston koko 324 kuukautta.

Python-versio 3.9.13, Spyder IDE -versio 5.2.2, TensorFlow-versio 2.8.0

```
#Moduulien ja pakettien lataaminen
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dropout
from tensorflow.keras.metrics import RootMeanSquaredError
from tensorflow.keras.optimizers import Adam

#Datan lataaminen, esikäsittely ja skaalaaminen
df = pd.read_excel("S:\YHTEISET\Omat kansiot\Henu\dttyodata2.xlsx")
dataset_orig=df.to_numpy()
scaler = MinMaxScaler(feature_range=(0, 1))
dataset = scaler.fit_transform(dataset_orig.reshape(-1,2))
df_transposed = dataset.T

# Asetetaan aiempien kuukausien ennuste LSTM-mallille
prediction=12
X=[]
y=[]
for i in range(prediction,len(dataset)):
    X.append(df_transposed[:,i-prediction:i])
    y.append(df_transposed[:,i])

X=np.array(X)

#Muodostetaan oppimis- ja testausaineistot
train_data_y=np.array(y)
train_data_y2=train_data_y.reshape(2,-1)
test_start = int(0.7*len(dataset_orig))
trainX=X[:test_start]
testX=X[test_start:]
trainY=train_data_y[:test_start]
testY=train_data_y[test_start:]

tf.random.set_seed(7)

#Luodaan ja sovitetaan LSTM-neuroverkko
model = Sequential()
model.add(LSTM(512, activation='relu',
input_shape=(2,(prediction)),return_sequences=True))
model.add(Dropout(0.2))
model.add(LSTM(256, activation='tanh',
```

```

input_shape=(2,(prediction)),return_sequences=True))
model.add(Dropout(0.2))
model.add(LSTM(256, activation='tanh',
input_shape=(2,(prediction)),return_sequences=False))
model.add(Dropout(0.2))
model.add(Dense(128, activation='relu'))
model.add(Dense(10, activation='tanh'))
model.add(Dense(2, activation='relu'))

#LSTM-neuroverkon koonti
optimizer = Adam(learning_rate=0.001)
model.compile(loss='mean_squared_error', optimizer=optimizer,
metrics=[RootMeanSquaredError()])
print(model.summary())
history = model.fit(trainX, trainY, epochs=1000, batch_size=4, verbose=1)

#Muodostetaan ennusteet
vektori=trainX[-1:]
lista=[]
for i in range(test_start+prediction,len(dataset)):
    predicted = model.predict(vektori)
    lista.append(predicted)
    predicted = np.expand_dims(predicted.T, axis=0)
    vektori=np.concatenate((vektori[-2:],predicted),axis=2)
    vektori=np.delete(vektori,0,axis=2)

lista_np1=np.array(lista)
lista_np2=np.squeeze(lista_np1)

#Käänteismuunnos
predicted_y = scaler.inverse_transform(lista_np2)
real_y=dataset_orig[test_start+prediction:]

#Muodostetaan RMSE ja MSE
plt.figure(figsize=(10, 6))
plt.plot(history.history['root_mean_squared_error'], label='RMSE')
plt.plot(history.history['loss'], label='MSE')
plt.xlabel("Laskentakierrokset")
plt.ylabel("Estimaatin arvo")
plt.legend()
plt.show()

#Verrataan opetusaineistoa todelliseen aineistoon
y_data = np.column_stack((real_y, predicted_y))
plot_data1 = y_data[:, [0]]
plot_data2 = y_data[:, [2]]
plt.plot(plot_data1, label='Todelliset havainnot')
plt.plot(plot_data2, label='Mallinnetut havainnot')
plt.xlabel("Aikapisteet")
plt.ylabel("Vahinkojen lukumäärä")
plt.legend()

#Mallin RMSE
testScoreLSTM = np.sqrt(mean_squared_error(real_y[:, 0], predicted_y[:, 0]))
print('Test ScoreLSTM: %.2f RMSE' % (testScoreLSTM))

#Muodostetaan 12kk ennuste
n_future = 12
predictions_future = model.predict(trainX[-n_future:])

```

```
y_pred_future = scaler.inverse_transform(predictions_future)

#Ennusteen lisääminen alkuperäisen aikasarjaan
test_plot = np.ones_like(dataset_orig) * np.nan
yhd = np.concatenate((test_plot[:, [0]],y_pred_future[:, [0]]))
plt.figure(figsize=(10, 6))
plt.plot(dataset_orig[:, [0]], label="Vahingot")
plt.plot(yhd[:, [0]], label="Ennuste")
plt.xlabel("Kuukaudet")
plt.ylabel("Vahinkojen lukumäärä")
plt.legend()
```

LIITE D: GRU-NEUROVERKKO

GRU-neuroverkko rakennusteollisuuden työtaturmien ennustamiseen. Aineiston koko 324 kuukautta.

Python-versio 3.9.13, Spyder IDE -versio 5.2.2, TensorFlow-versio 2.8.0

```
#Moduulien ja pakettien lataaminen
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dropout
from tensorflow.keras.metrics import RootMeanSquaredError
from tensorflow.keras.optimizers import Adam

#Datan lataaminen, esikäsittely ja skaalaaminen
df = pd.read_excel("S:\YHTEISET\Omat kansiot\Henu\dtyodata2.xlsx")
dataset_orig=df.to_numpy()
scaler = MinMaxScaler(feature_range=(0, 1))
dataset = scaler.fit_transform(dataset_orig.reshape(-1,2))
df_transposed = dataset.T

# Asetetaan aiempien kuukausien ennuste LSTM-mallille
prediction=12
X=[]
y=[]
for i in range(prediction,len(dataset)):
    X.append(df_transposed[:,i-prediction:i])
    y.append(df_transposed[:,i])

X=np.array(X)

#Muodostetaan oppimis- ja testausaineistot
train_data_y=np.array(y)
train_data_y2=train_data_y.reshape(2,-1)
test_start = int(0.7*len(dataset_orig))
trainX=X[:test_start]
testX=X[test_start:]
trainY=train_data_y[:test_start]
testY=train_data_y[test_start:]

tf.random.set_seed(7)

#Luodaan ja sovitetaan LSTM-neuroverkko
model = Sequential()
model.add(LSTM(512, activation='tanh',
input_shape=(2,(prediction)),return_sequences=True))
model.add(Dropout(0.2))
```

```

model.add(GRU(128, activation='tanh',
input_shape=(2,(prediction)),return_sequences=True))
model.add(Dropout(0.2))
model.add(GRU(128, activation='tanh',
input_shape=(2,(prediction)),return_sequences=False))
model.add(Dropout(0.2))
model.add(Dense(10, activation='tanh'))
model.add(Dense(10, activation='tanh'))
model.add(Dense(2, activation='relu'))

#LSTM-neuroverkon koonti
optimizer = Adam(learning_rate=0.001)
model.compile(loss='mean_squared_error', optimizer=optimizer,
metrics=[RootMeanSquaredError()])
print(model.summary())
history = model.fit(trainX, trainY, epochs=1000, batch_size=4, verbose=1)

#Muodostetaan ennusteet
vektori=trainX[-1:]
lista=[]
for i in range(test_start+prediction,len(dataset)):
    predicted = model.predict(vektori)
    lista.append(predicted)
    predicted = np.expand_dims(predicted.T, axis=0)
    vektori=np.concatenate((vektori[-2:],predicted),axis=2)
    vektori=np.delete(vektori,0,axis=2)

lista_np1=np.array(lista)
lista_np2=np.squeeze(lista_np1)

#Käänteismuunnos
predicted_y = scaler.inverse_transform(lista_np2)
real_y=dataset_orig[test_start+prediction:]

#Muodostetaan RMSE ja MSE
plt.figure(figsize=(10, 6))
plt.plot(history.history['root_mean_squared_error'], label='RMSE')
plt.plot(history.history['loss'], label='MSE')
plt.xlabel("Laskentakierrokset")
plt.ylabel("Estimaatin arvo")
plt.legend()
plt.show()

#Verrataan opetusaineistoa todelliseen aineistoon
y_data = np.column_stack((real_y, predicted_y))
plot_data1 = y_data[:, [0]]
plot_data2 = y_data[:, [2]]
plt.plot(plot_data1, label='Todelliset havainnot')
plt.plot(plot_data2, label='Mallinnetut havainnot')
plt.xlabel("Aikapisteet")
plt.ylabel("Vahinkojen lukumäärä")
plt.legend()

#Mallin RMSE
testScoreGRU = np.sqrt(mean_squared_error(real_y[:, 0], predicted_y[:, 0]))
print('Test ScoreGRU: %.2f RMSE' % (testScoreGRU))

#Muodostetaan 12kk ennuste
n_future = 12

```

```
predictions_future = model.predict(trainX[-n_future:])
y_pred_future = scaler.inverse_transform(predictions_future)

#Ennusteen lisääminen alkuperäisen aikasarjaan
test_plot = np.ones_like(dataset_orig) * np.nan
yhd = np.concatenate((test_plot[:, [0]],y_pred_future[:, [0]]))
plt.figure(figsize=(10, 6))
plt.plot(dataset_orig[:, [0]], label="Vahingot")
plt.plot(yhd[:, [0]], label="Ennuste")
plt.xlabel("Kuukaudet")
plt.ylabel("Vahinkojen lukumäärä")
plt.legend()
```