

**Bachelor in Medien- und Kommunikationsinformatik
SAT**

WS2018

Prof. Dr. Christian Kücherer

Erstfassung

**Wie können neuronale Netze in der
Softwaretechnik eingesetzt werden und wo
sind die aktuell bekannten Grenzen**

vorgelegt von:

**Konstantin
Rosenberg**

MKI, 6. Semester

**Robin
Schramm**

MKI, 6. Semester

**Ravell
Heerdegen**

MKI, 6. Semester

Kontaktadressen:

konstantin_nils.rosenberg
@Student.reutlingen-
university.de

robin_connor.schramm
@Student.reutlingen-
university.de

ravell.heerdegen
@Student.reutlingen-
university.de

Eingereicht am: 20.11.2018

Abstract

Künstliche neuronale Netze haben in den letzten zwanzig Jahren ein exorbitantes Interesse für vielerlei Bereiche erweckt. Ihre Entwicklung geht bis in die fünfziger Jahre zurück und erfährt seitdem immer wieder neue Fortschritte. Dies liegt vor allem daran, dass sie sich weiterentwickeln und dazulernen können und somit auch für zukünftige Themen und Aufgaben geeignet sind. Besonders im Bereich der Softwaretechnik sind künstliche neuronale Netze heute nicht mehr wegzudenken. In dieser Arbeit wird näher auf die Disziplinen Mustererkennung, Kosten- und Aufwandsschätzung sowie Qualitätsmanagement eingegangen und Beispiele aufgezeigt, um die Frage zu beantworten, wie neuronale Netze in der Softwaretechnik eingesetzt werden können und wo die aktuell bekannten Grenzen sind. Dafür werden relevante Begriffe und Methoden erläutert sowie handfeste Beweise referenziert. Auch vergleichen wir in unserer Arbeit die verschiedenen Arten von neuronalen Netzen und zeigen Vor- sowie Nachteile auf. Das Ergebnis unserer Arbeit ist, dass sich neuronale Netze fest in der Softwaretechnik etabliert haben, und jede aufgeführte Art in jeder der beschriebenen Disziplinen sowohl eindeutige Verbesserungen als auch Defizite gegenüber klassischen Methoden aufweist. Daraus resultiert, dass es noch Verbesserungsbedarf sowie -möglichkeiten für die Entwicklung von künstlichen neuronalen Netzen gibt. In Zukunft könnten sich noch mehr Gebiete erschließen, welche mit neuronalen Netzen gekoppelt werden, um schnellere und bessere Ergebnisse zu erzielen. Vorstellbar ist auch, dass sich komplett neue Arten von neuronalen Netzen bilden, da selbst heute schon diverse Typen kombiniert werden, um zu neuen Erkenntnissen zu gelangen.

Contents

| | | |
|----------|---|-----------|
| 1 | Artificial neural networks | 3 |
| 2 | Convolutional neural networks | 3 |
| 3 | Residual neural networks | 3 |
| 4 | Back propagation | 4 |
| 5 | Hidden Markov Model | 4 |
| 6 | Pattern recognition | 4 |
| 6.1 | Speech recognition | 6 |
| 6.1.1 | Speech recognition und convolutional neural networks . | 6 |
| 6.1.2 | Speech recognition und residual neural networks | 7 |
| 6.2 | Emotion recognition | 8 |
| 6.2.1 | Emotion recognition und convolutional neural networks | 8 |
| 6.3 | Fazit | 10 |
| 7 | Cost und effort estimation | 11 |
| 7.1 | Stand der Technik | 11 |
| 7.2 | COCOMO | 12 |
| 7.3 | Soft Computing Ansätze | 13 |
| 7.4 | Leistungsbeurteilung | 13 |
| 7.5 | Vergleich | 13 |
| 7.6 | Fazit | 14 |

1 Artificial neural networks

Artificial neural networks (ANN) basieren auf der Funktion des biologischen Nervensystems. Informationen werden durch eine Vielzahl miteinander verknüpfter parallel laufender Einheiten verarbeitet. Die daraus entstehenden Systeme trainieren sich teilweise selbst, um schnellere Entscheidungen treffen und Fehler vermeiden zu können. Potenzielle Systeme für Speech recognition sind u.a. sogenannte Single-layers, perception models, multi-layer perception model oder predictive neural networks.[1]

2 Convolutional neural networks

Convolutional neural networks (CNNs) sind sogenannte feed-forward neural networks. Das bedeutet, dass Signale durch die verschiedenen Layer (Schicht) des neuronalen Netzes (NN) gegeben werden. CNNs bestehen aus input layers, convolutional layers, pooling layers, sowie fully-connected layers und output layers.[2] Jeder Layer besteht aus einer Vielzahl an zweidimensionalen Flächen. Eine Fläche besitzt mehrere unabhängige Neuronen. Eingegebene Spracheingaben werden in eine bestimmte Menge an Frequenzen unterteilt. Jede einzelne Frequenz wird mit sogenannten Gewichten verglichen, welche mit einer lokalen Variable im Layer multipliziert werden. Die Gewichte werden auf den gesamten Eingaberaum verteilt. Die Verteilung aktiviert die convolutional layers und teilt diese in verschiedene Bereiche auf. Jeder convolutional layer besitzt diverse Filter. Durch die Filter entstehen sogenannte hidden units. Hidden units sind beschreibbare Plätze auf einem Layer. Im Anschluss werden die hidden units in einen max-pooling layer gegeben. Im max-pooling layer werden Varianzen der einzelnen hidden units entfernt. Die Varianzen können u.a. aus Dialekten, undeutlicher Sprache oder Störungen in der Übertragung kommen. Die nun bereinigten hidden units werden in den convolutional layers als Trainingsparameter eingespeichert. Somit erkennt jeder convolutional layer verschiedene Merkmale, aufgrund der ihm zur Verfügung stehenden Merkmalsammlung.[3]

3 Residual neural networks

Residual neural networks (Resnets) gehören zur Kategorie der deep neural networks (DNNs). Resnets bestehen aus sogenannten residual blocks (Res). Ein residual block besteht aus weight layers sowie einem dropout, auch dropout regularization genannt. [4] Ein Layer besitzt mehrere units, also Neuronen, welche trainiert werden können. Der dropout sorgt dafür,

dass Neuronen nicht überangepasst werden. Das geschieht durch zufälliges Ausschalten bestimmter Neuronen, was durch eine dropout rate festgelegt wird. Dadurch müssen andere Neuronen die Arbeit des Neurons erledigen, welches zuerst für die Aufgabe vorgesehen war. Eine Überanpassung eines Neurons hat zur Folge, dass Trainingsdaten nicht mehr aufgrund ihrer Eigenschaften, sondern aufgrund des vorangegangenen Trainings erkannt werden. In [4] wird sowohl die Architektur als auch die Funktionalität eines Resnet genauer beschrieben.

4 Back propagation

Multilayer perceptrons (mehrschichtige neuronale Netze) benutzen u.a. Back propagation, um Parameter für künstliche neuronale Netze zu trainieren. Back propagation dient dem Trainieren von Layern und ihren feature maps, um auf entstandene Fehler aufmerksam zu machen, welche rückwärts durch das NN gegeben werden.[3]

5 Hidden Markov Model

Das Hidden Markov Model (HMM) ist bekannt für seinen Nutzen im Bereich der speech recognition. HMM ist in der Lage Wörter zu modellieren, welche daraufhin einem bestimmten Kontext zugeordnet werden können. Das HMM besteht aus zwei Teilen. Der sogenannte hidden part besteht aus Zuständen, wobei jeder einzelne einem Phonem entspricht. Der observable part dient der statistischen Präsentation der akustischen Gegebenheiten. Der HMM-Algorithmus gehört zu den unsupervised learning techniques, da er Parameter unkontrolliert schätzt ohne zusätzliche Kennzeichnungen durch Menschen zu benötigen. Auch kann das HMM zur Dekodierung verwendet werden. Wenn ein einzelnes HMM benutzt wird, bedeutet Dekodierung, eine bestimmte Reihe an Zuständen zu finden und jede einzelne Eingabe einem Zustand zuzuordnen.[5]

6 Pattern recognition

Pattern recognition zeichnet sich durch die Möglichkeit aus, Klassifizierungen durchzuführen, welche entweder überwacht oder unüberwacht, also automatisch und ohne menschliche Hilfe ablaufen. Im Bereich pattern recognition haben sich in den letzten Jahren besonders neuronale Netze als enormer Fortschritt herausgestellt. Mit neuronalen Netzen werden u.a. Muster erkannt

und zugeordnet, Umgebungen analysiert, z.B. im Bereich emotion recognition, Merkmale erkannt und spezifischen Klassen zugeordnet oder auch Bewertungen von Performanz durchgeführt.[6] Im folgenden Kapitel werden zwei Disziplinen bezüglich pattern recognition beschrieben. Der erste Abschnitt befasst sich mit der Disziplin speech recognition. Dabei gehe ich auf die Definition von speech recognition ein und zeige anschließend diverse Beispiele für den Einsatz von neuronalen Netzen in selbiger Disziplin auf. Im zweiten Abschnitt folgt die Disziplin emotion recognition sowie einige Beispiele für den Einsatz von neuronalen Netzen in der selbigen Disziplin. Abschließend folgt ein Fazit, welches den Inhalt des Kapitels kurz zusammenfasst und reflektiert.

6.1 Speech recognition

Speech recognition, auch automatic speech recognition (ASR) genannt, macht es Maschinen möglich, Sprache in Text oder Kommandos zu übersetzen. Dies geschieht sowohl durch Identifikation als auch durch Verstehen der eingegebenen Sprachsignale. Die Sprachsignale werden auf erlernte Muster angewendet, wobei Merkmale entnommen werden, um die Signale zuordnen zu können. Das Ziel ist, die menschliche Sprache zu verstehen, um mit dem Menschen kommunizieren zu können.[1]

6.1.1 Speech recognition und convolutional neural networks

In [3] beschreibt Du Guiming et. al. einen Versuch, bei dem Speech recognition in Verbindung mit einem CNN getestet wird. Die dafür verwendeten Trainingsdaten stammen von 30 Personen. Für die automatische Erkennung der Sprachsignale und ihrer Tonhöhen wird der Mel Frequency Cepstral Coefficient (MFCC) eingesetzt, welcher das Frequenzspektrum kompakt darstellen kann. Für die Merkmalentnahme kommen hamming windows zum Einsatz. Mit dieser Technik können Signale gefiltert und Varianzen geglättet werden, bevor sie das neuronale Netz erreichen. Nach der Verarbeitung der Signale, wird das neuronale Netz mit fünf Personen getestet. Beim ersten Durchlauf der Signale durch das neuronale Netz, sind die erwarteten und vorhergesehenen Werte alle gleich. Der Unterschied zwischen erwarteten und echten Werten wird als cost function bezeichnet. Ist die cost function niedrig, sind die Parameter hochwertiger. Es zeigt sich, dass mit steigender Durchlaufzahl die cost function niedriger wird. Somit kann bestätigt werden, dass sich CNNs sowohl für die Erkennung von isolierten Wörtern eignet, als auch für die Reduzierung von Varianzen.[3]

Rafael M. Santos et. al. zeigen in [7] auf, wie sich ein CNN in einer geräuschvollen Umgebung verhält. Da sich die sogenannten shared weights in einem CNN über den gesamten Eingaberaum erstrecken, können Merkmale selbst dann gefunden werden, wenn sie sich aufgrund von Störungen z.B. durch Lautstärke an unterschiedlichen Positionen befinden. In dem Experiment wird das CNN so trainiert, dass es die HMM-Verteilungswahrscheinlichkeit schätzen kann. Außerdem klassifiziert das CNN nicht nur die aktuelle Aufnahme, sondern bezieht vorherige und nachkommende Aufnahmen mit ein, um Zusammenhänge zwischen den Informationen herstellen zu können. Als Eingabe werden einzelne Wörter mit ansteigender Lautstärke eingesetzt. Die Eingabedaten bestehen aus brasilianischen und portugiesischen Wörtern,

welche zehn mal von jeweils sechs Männern und zwei Frauen vorgetragen wurden. Die Aufnahmen wurden in einer unkontrollierten Umgebung mit einem Mobiltelefon aufgezeichnet, von welchem 8000 Samples in der Sekunde entnommen wurden. Bei dem Versuch hat sich ergeben, dass die Kombination von CNN und HMM, sowohl feature extraction als auch signal detection durchführt. Auch wurde herausgefunden, dass sich die Kombination CNN-HMM an wenigsten von Lautstärke und Geräuschen beeinflussen lässt, im Vergleich zu anderen getesteten Kombinationen wie GMM oder SVM [8] in Verbindung mit HMM.

6.1.2 Speech recognition und residual neural networks

In der Studie [4] wurde ein Resnet mit mehreren Res in Kombination mit Speech recognition getestet. Es hat sich zunächst herausgestellt, dass mit einer Anzahl von acht Res, welche hintereinander geschaltet werden, die beste Performanz in Bezug auf die word error rate (WER) erreicht wird. Ab einem Wert von mehr als acht Res, hat die Performanz abgenommen. Verglichen wurde das sogenannte 8Res dabei mit einem 8-DNN, welches aus sechs hidden-layers bestand. Dabei hat sich ergeben, dass das 8Res bessere Ergebnisse als das 8-DNN erzielt. Auch wird aufgezeigt, wie sich wide residual networks (WResnets) auf die Validierung von Daten auswirkt. Wurde die Anzahl der Layer im getesteten Resnet von vier auf zehn erhöht, wurde eine noch bessere Performanz erreicht. Ab einem Wert von zwölf nahm die Performanz wiederum ab, da die Neuronen ab diesem Wert überangepasst waren. Schließlich wurde festgestellt, dass HMM-Resnets mit steigender Anzahl an Res und Layern eine um 0,4% bessere Performanz als ein HMM-DNN in Bezug auf die WER erreicht haben.

6.2 Emotion recognition

Emotion recognition, auch automatic emotion recognition (AER) genannt, ist ein interdisziplinäres Forschungsgebiet, welches u.a. die Bereiche Computer Sciences, Cognitive Science sowie Psychology vereint. Das Erforschen von emotionalen Zuständen ist hilfreich, um menschliche Aktionen zu verstehen, als auch um menschliche Faktoren in künstliche Systeme zu integrieren.[9] Generell wird das Erfassen von Emotionen über die Muskelbewegungen im menschlichen Gesicht geregelt. Die Muskelbewegungen im Gesicht werden auch action units (AUs) genannt. Nachdem ein Gesicht erfasst und isoliert wurde, kann man dem Gesicht eine Emotion zuordnen mithilfe von action units analysis.[10] Beim Arbeiten im Bereich Emotion recognition und facial expression recognition, unterscheidet man zwischen sechs generellen, kultur-unabhängigen Emotionen, welche 1971 von Ekman und Friesen als universell festgelegt wurden: Wut, Ekel, Angst, Freude, Trauer und Überraschung.[11] Emotion recognition findet aber auch Anwendung in Verbindung mit Sprache, bzw. Körpersprache. Mit Emotion recognition möchte man natürliche Kommunikation zwischen Mensch und Maschine herstellen und verbessern, jedoch auch Zustände eines Menschen erkennen, um anschließend passende Werbung oder zugeschnittene Angebote zu erstellen. Methodisch wird zwischen static approaches und dynamic approaches unterschieden. Das bedeutet, einerseits mit statischen Bildern zu arbeiten, andererseits mit Sequenzen von Bildern, u.a. auch Videos. Emotion expression recognition wird grundsätzlich in die Schritte face image acquisition, feature extraction und facial emotion expression recognition unterteilt. [2]

6.2.1 Emotion recognition und convolutional neural networks

In [2] wird ein Versuch mit einem CNN in Verbindung mit automatic facial expression recognition (AFER) beschrieben. Bei dem Versuch wurde das Facial Action Coding System (FACS) verwendet, um die Genauigkeit der Erkennung von Emotionen zu verbessern. Das geschieht aufgrund der Möglichkeit des FACS, die AUs im Gesicht zu erkennen und zu interpretieren. Auch kann mit dem FACS die Intensität einer Emotion definiert werden. Als Trainingsdaten fungierte der Cohn-Kanade (CK+) Datensatz. Der CK+ besteht aus 123 Individuen, welche verschiedene Folgen von emotionalen Ausdrücken aufgenommen haben. Die daraus entstandenen Bilder wurden auf eine einheitliche Größe skaliert und verpixelt. Zusätzlich kamen pre-processing techniques wie z.B. Rotations- oder Farbkorrektur zum Einsatz, bevor die Daten in das CNN gegeben wurden. Im ersten convolutional layer

des eingesetzten CNN wurden visuelle Merkmale wie Kanten, Lippenformen, Falten sowie auch Augen und Augenbrauen extrahiert. Durch die berechneten shared weights zwischen den Schichten im CNN, kann ein hohes oder ein geringes Vorkommen von AUs im output layer festgestellt werden, was auf verschiedene Klassifizierungen von Emotionen schließen lässt. Ergebnis des Versuches war, dass das trainierte CNN Parallelen zwischen den Emotionen Wut und Neutralität gefunden hat, was mit den ähnlichen Gesichtszügen von Wut und Neutralität zusammenhängt. Auch hat sich ergeben, dass Freude am besten erkannt wurde. Dadurch wurde bestätigt, dass sich CNNs für die Erkennung von Emotionen in einem menschlichen Gesicht eignen.

Das Experiment [10] benutzt ebenfalls ein CNN, jedoch mit anderen Techniken. Zum Einsatz kommen zwei Methoden, welche sich für das Erfassen von sogenannten lokalen sowie globalen Informationen eignen. Die eine, genannt bottom-up, eignet sich zum Isolieren von Gesichtern aus Bildern, um diese anschließend in das bereits trainierte CNN als Input einzugeben. Die andere nennt sich top-down und bezeichnet sich als labeling algorithm. Beim labeling werden label, auch descriptors genannt, einer Szene zugeordnet, um den Kontext der Szene zu erfassen. Die descriptors werden daraufhin in ein bayesian network (BN) gegeben, welches die descriptors verarbeitet, um Beziehungen und Abhängigkeiten zwischen den Begriffen herzustellen. Die verarbeiteten descriptors werden als Input in das CNN geleitet. Der output layer des CNN besteht dadurch aus drei Ausgängen, welche die Szene als negativ, positiv oder neutral einordnen. Zum Einsatz kommt außerdem die GAF Datenbank, welche 6470 Bilder beinhaltet. Diese Menge wurde aufgeteilt auf die Trainingsdaten, Validierungsdaten und auf die Testdaten. Zusätzlich wurde ein dropout eingesetzt, um die Überanpassung der Neuronen zu vermeiden. Endlich hat sich ergeben, dass das CNN in Kombination mit einem BN die besten Ergebnisse liefert, im Gegensatz zum Versuch ohne BN bzw. ohne CNN. Ein interessanter Ausbau des Versuches wird am Ende gegeben. Dabei wird erwähnt, dass die Klassifizierungen negativ, positiv und neutral, auf drei CNNs aufgeteilt werden könnte, um jedes zu spezialisieren und somit noch bessere Ergebnisse erzielen zu können.

6.3 Fazit

Vergleicht man die aufgeführten Beispiele und dessen Ergebnisse miteinander, so kommt man zu dem Schluss, dass neuronale Netze durchaus Anwendung im Bereich pattern recognition finden, was wiederum belegt, dass neuronale Netze im Bereich Softwaretechnik eingesetzt werden. Auch wurde einerseits gezeigt, dass sowohl CNNs als auch Resnets für Speech recognition eingesetzt werden. Andererseits wurde aufgezeigt, dass sich CNNs auch im Bereich emotion recognition etabliert haben und wie in [10] bewiesen, sogar am Besten für facial emotion expression recognition in unkontrollierten Umgebungen eignen. Ebenfalls wurden verschiedene Ansätze für den Einsatz von CNNs gegeben, was zeigt, dass diese auf unterschiedliche Weise fungieren können, um diverse Aufgaben zu erledigen. Verfolgt man die angegebenen Referenzen dieses Kapitels, so erhält man eine Vielzahl ähnlicher Versuche und Experimente, welche die Flexibilität, Eignung und Einsatzmöglichkeiten von NNs im Bereich pattern recognition nochmals bestätigen und unterstreichen. Auch wird bezogen auf die aufgeführten Beispiele deutlich, dass die genannten CNNs mitunter zu den momentan besten Lösungen für die Aufgaben Speech- sowie emotion recognition gehören. Erwähnenswert ist außerdem, dass unter aktuelle Grenzen u.a. die Fehlerrate für das Erkennen von Sprachsignalen und Emotionen gehört, welche momentan noch keine einhundert prozentige Zuordnung erreicht, Nebeneffekte wie die Überanpassung von Neuronen aufgrund zu tiefer Spezialisierung, das Zusammenführen von Wut und Neutralität in der Disziplin emotion recognition aufgrund der nicht zu identifizierbaren Unterschiede, sowie die Technik, welche die NNs antreibt und somit den Ausbau der selbigen einschränkt.

7 Cost und effort estimation

In den sechziger und siebziger Jahren stieg die Größe und Komplexität von Softwareprojekten drastisch an. Softwareentwicklung wurde immer mehr kommerziell genutzt wodurch sich neue Arten von Kunden bildeten. Darunter vor allem das Militär. Durch die Vergrößerung der Branche war dies immer weniger effektiv ohne vorherige Planung zu entwickeln. Als Lösung für das Problem entstand das Software Projekt Management(SPM), welches auch heutzutage noch eine zentrale Rolle in der Softwareentwicklung spielt.[12] Einer der Punkte mit dem sich SPM beschäftigt ist das project cost management(PCM). Nach Bajita[12] unterteilt sich PCM in die Kategorien software cost estimation(SCE) und software effort estimation(SEE). Aktuelle Forschungen zeigen, dass genaue SCE und SEE die Chance auf ein qualitativ hochwertiges Produkt erhöht.[13][14] Jeffery et. al. beschreiben, dass SCE eines der wichtigen Werkzeuge ist, um einen kompetitiven Vorteil im Bereich des IT-Service zu erhalten.[15] Aus dieser Notwendigkeit entstanden viele SCE und SEE Tools, welche Experten bei der Projektplanung unterstützen sollen.

Im Zuge dieses Kapitels werden als Erstes die bestehenden SCE und SEE Techniken vorgestellt und kategorisiert. Als nächstes wird das klassische constructive cost estimation model(COCOMO) sowie neuere Soft Computing Modelle vorgestellt und deren Vorteile und Probleme aufgelistet. In den weiteren Absätzen wird erklärt, wie die Ergebnisse der Schätzungen gemessen und in Zahlen gefasst werden können, um danach auf einige Beispiele von direkten Vergleichen der mathematischen und Soft Computing einzugehen. Zum Schluss wird anhand der vorherigen Abschnitte beantwortet, ob NNs in der Softwaretechnik(SWT) im Bereich der SCE und SEE eingesetzt werden können und wo deren derzeit bekannte Grenzen liegen.

7.1 Stand der Technik

In der Planungsphase eines Softwareprojekts ist es essentiell, die Kosten und die Dauer des Projekts zu kennen. Kunden geben Aufträge und müssen wissen, wie viel welche Funktion kosten würde und wann das Projekt voraussichtlich abgeschlossen sein wird. Grobe Schätzungen können allein durch Erfahrung der Entwickler abgegeben werden, wie zum Beispiel in SCRUM, bei dem User Stories je nach Komplexität Story Points gegeben werden. Bei ansteigender Komplexität Projekten kann die manuelle Schätzung nicht ausreichen. Das zeigt sich im großen Zweig des SPM, welcher sich allein mit der Planung und Schätzung von Softwareprojekten auseinandersetzt. Eine wichtige Metrik in der SPM ist die Größe der Software. Dabei kann man zwei

Arten von Methoden der Größenbestimmung unterscheiden. Zum Einen die gezählten Lines of Code (LOC) und die daraus abgeleiteten 1000 Lines of Code (KLOC). Zum anderen die function points (FPs). Anhand der Arten der Modellkonstruktion lassen sich SCE und SEE Modelle in vier Kategorien unterteilen. Erstens Expertensysteme, die sich auf die Erfahrung der Personen verlassen[16]. Zweitens die Linearen Systeme, die auf der Größenmessung mit FP basieren[13]. Drittens die nonlinearen Mathematischen Modelle. Viertens die neueren Soft Computing Ansätze, die entweder bisherige Methoden erweitern oder komplett neu aufbauen[17].[18]. Eines der beliebtesten Tools ist COCOMO.[19] COCOMO lässt sich den nichtlinearen Systemen zuordnen.

7.2 COCOMO

COCOMO wurde 1981 von Boehm entworfen, um Fehler im SPM zu verringern. COCOMO ist ein algorithmisches Modell mit vielen Eingabeparametern(zahl). In COCOMO gibt es drei Typen von Projekten, gemessen an der Komplexität. Organic Mode für einfache kleine Projekte, embedded mode für innovative komplexe Projekte mit hohen Anforderungen sowie den semi-detached mode für Projekte zwischen dem organic und embedded type. Der Projekttyp wirkt sich neben den LOC signifikant auf die Endschätzung aus. Zudem nutzt COCOMO einige cost drivers, welche je nach Art variieren. In Boehms ersten COCOMO gab es 15 cost drivers. Cost drivers werden mit Werten zwischen *sehr niedrig* und *sehr hoch* bewertet. Die Werte stellen dabei Gewichtungen dar, die sich ebenfalls auf den Endwert auswirken. Das Ergebnis stellt die Schätzung der Kosten in einer gewünschten Währung dar, sowie die benötigten Personenmonate. Die klassischen Modelle bringen Probleme mit sich, von denen hier einige beschrieben werden.

- Algorithmische Tools sind generisch und somit nicht an das spezifische Projekt angepasst, daraus ergibt sich ein Mangel an bestehenden Daten für die eigene Nische. Daten von anderen Firmen sind in der Regel nicht verfügbar.
- Erfahrung in großen Projekten ist schwierig, da diese eine lange Lebensdauer haben, bis sie abgeschlossen werden und die Schätzung mit dem Ergebnis verglichen werden kann.
- Oft werden die Schätzungen nebenbei gemacht und nicht als eigenständiger Punkt in der Planung beachtet, was die Wahl der Eingabeparameter negativ beeinflusst. Die Eingabeparameter beeinflussen die Schätzung jedoch maßgeblich
- Mangel an Flexibilität bei sich ändernden Anforderungen. Besonders im Agilen Bereich.
- Schwierigkeit und Aufwand, Projekt spezifische Daten zu sammeln, zu

beurteilen und in konkreten Zahlen auszudrücken. [20][16][21]
Einige der genannten Probleme versuchen verschiedene Forscher mit Soft Computing Ansätzen zu lösen[14]

7.3 Soft Computing Ansätze

Unter Soft Computing versteht man vorallem Fuzzy Logic, Evolutionäre Algorithmen und Künstliche Neurale Netze (NNs). Mit ihnen kann das Problem gelöst werden, nur an ein Projekt angepasst zu sein, da sie sich leicht generalisieren lassen. Zudem sind sehr flexibel und anpassungsfähig, was sie bei sich ändernden Anforderungen wertvoll macht.[22] Damit sie zufriedenstellende Ergebnisse liefern, müssen die Trainingsdaten passend gewählt sein. NNs mit zufällig gewählten Trainingsdaten liefern keine besseren Schätzungen als Regressionsmodelle. [23]

7.4 Leistungsbeurteilung

Um die Modelle untereinander vergleichen zu können, gibt es verschiedene Methoden die Qualität zu messen. Nach Finnie et. al. ist der mean absolute relative error (MARE) die bevorzugte Fehlerbestimmungsmethode von Software-Messforschern. Der MARE wird wie folgt berechnet:

$$\left| \frac{estimate - actual}{estimate} \right| \quad (1)$$

estimate ist dabei der vom Modell geschätzte Wert und *actual* der nach der Durchführung tatsächlich gemessene Wert. Um herauszufinden, ob das Modell über- oder unterschätzt, wird der mean relative error (MRE) wie folgt berechnet:

$$\left(\frac{estimate - actual}{estimate} \right) * 100 \quad (2)$$

Ein Wert von Null ist dabei ein perfektes Ergebnis. Je höher oder niedriger die Zahl, desto größer wurde über- oder unterschätzt.[24] Ein Ergebnis zwischen -25 und 25 gilt als gutes Ergebnis in SEE.[21]. Viele der vorgestellten Modelle werden an laufenden Industrieprojekten getestet. Das am weitesten akzeptierte Evaluationskriterium ist der MRE aus einem Satz von 63 NASA Projekten.[25]

7.5 Vergleich

Abrahamsson et. al. vergleichen in ihrem Artikel inkrementelle Ansätze mit globalen Ansätzen im Kontext der agilen Entwicklung. Die verwen-

ten Modelle sind einerseits die Schrittweise lineare Regression und andererseits zwei verschiedene Arten von NNs. Eines ihrer Ergebnisse war, dass NNs nicht generell besser sind als einfache Regression. Das Multilayer Perceptron (MLP) aus dem Test liefert zum Beispiel schlechtere Ergebnisse als die meisten Regressionsmodelle. Das Radial Basis Function (RBF) NN lieferte ähnliche Ergebnisse wie das Regressionsmodell.[21] Besser als COCOMO[25] Auch besser als COCOMO in NASA Testprojekten in über 90

7.6 Fazit

Dieses Kapitel hat verschiedene Möglichkeiten aufgezeigt, wie NNs in den wichtigen Bereichen der SCE und SEE innerhalb der SWT eingesetzt werden. Wenn für gewisse Projekte passende NNs ausgewählt und korrekt konfiguriert werden, liefern NNs bessere Ergebnisse als reine mathematische Modelle. Vorallem in der agilen Entwicklung lösen NNs viele Probleme anderer Ansätze, auch in Verbindung und als Erweiterung der anderen Modelle. Dennoch sind sie nicht in ausschließlich allen Fällen die beste Lösung. Es existiert noch kein NN, welches eine Universallösung für alle Softwareschätzungen darstellt. Für manche Projekte bietet sich die Lösung mit NNs mehr an als für andere, vorallem da die optimale Lösung noch nicht während der Planung bekannt ist. Vorallem in Umgebungen mit wenigen Daten oder wenig Zeit sind NNs noch nicht praktikabel, da ein passendes NN zuerst entwickelt und trainiert werden muss.

References

- [1] Meng, J., Zhang, J., Zhao, H.: Overview of the speech recognition technology. Proceedings - 4th International Conference on Computational and Information Sciences, ICCIS 2012 pp. 199–202 (2012). <https://doi.org/10.1109/ICCIS.2012.202>
- [2] Awasthi, A.: Facial Emotion Recognition Using Deep Learning. IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI) Dec. 22, 2017 **1**(September), 9–12 (2013). <https://doi.org/10.1145/2818346.2830593>
- [3] Guiming, D., Xia, W., Guangyan, W., Yan, Z., Dan, L.: Speech recognition based on convolutional neural networks. 2016 IEEE International Conference on Signal and Image Processing, ICSIP 2016 pp. 708–711 (2017). <https://doi.org/10.1109/SIPROCESS.2016.7888355>
- [4] Vydana, H.K., Vuppala, A.K.: Residual neural networks for speech recognition. 25th European Signal Processing Conference, EUSIPCO 2017 **2017-Janua**, 543–547 (2017). <https://doi.org/10.23919/EUSIPCO.2017.8081266>
- [5] Tang, X.: Hybrid hidden markov model and artificial neural network for automatic speech recognition. Proceedings of the 2009 Pacific-Asia Conference on Circuits, Communications and System, PACCS 2009 pp. 682–685 (2009). <https://doi.org/10.1109/PACCS.2009.138>
- [6] A.~K.~Jain, R.~P.~W.~Duin, J.~Mao: {S}tatistical {P}attern {R}ecognition: {A} {R}eview. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 1, JANUARY 2000 **22**(1), 4–37 (2000)
- [7] Santos, R.M., Matos, L.N., Macedo, H.T., Montalvao, J.: Speech recognition in noisy environments with convolutional neural networks. Proceedings - 2015 Brazilian Conference on Intelligent Systems, BRACIS 2015 pp. 175–179 (2016). <https://doi.org/10.1109/BRACIS.2015.44>
- [8] Padmanabhan, J., Premkumar, M.J.J.: Machine learning in automatic speech recognition: A survey. IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India) **32**(4), 240–251 (2015). <https://doi.org/10.1080/02564602.2015.1010611>, <http://dx.doi.org/10.1080/02564602.2015.1010611>

- [9] Aracena, C., Basterrech, S., Snasel, V., Velasquez, J.: Neural Networks for Emotion Recognition Based on Eye Tracking Data. Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015 pp. 2632–2637 (2016). <https://doi.org/10.1109/SMC.2015.460>
- [10] Surace, L., Patacchiola, M., Sönmez, E.B., Spataro, W., Cangelosi, A.: Emotion Recognition in the Wild using Deep Neural Networks and Bayesian Classifiers. Proceeding ICMI 2017 Proceedings of the 19th ACM International Conference on Multimodal Interaction Pages 593-597 pp. 593–597 (2017). <https://doi.org/10.1145/3136755.3143015>, <http://arxiv.org/abs/1709.03820>
- [11] Huang, C.: Combining convolutional neural networks for emotion recognition. 2017 IEEE MIT Undergraduate Research Technology Conference, URTC 2017 **2018-Janua**, 1–4 (2018). <https://doi.org/10.1109/URTC.2017.8284175>
- [12] Bajta, M.E., Idri, A., Ros, J.N., Fernandez-Aleman, J.L., Gea, J.M.C.D., Garcia, F., Toval, A.: Software project management approaches for global software development: a systematic mapping study. Tsinghua Science and Technology **23**(6), 690–714 (2018). <https://doi.org/10.26599/TST.2018.9010029>
- [13] Matson, J.E., Barrett, B.E., Mellichamp, J.M.: Software development cost estimation using function points. IEEE Transactions on Software Engineering **20**(4), 275–287 (1994). <https://doi.org/10.1109/32.277575>
- [14] Bilgaiyan, S., Mishra, S., Das, M.: A Review of Software Cost Estimation in Agile Software Development Using Soft Computing Techniques. In: 2016 2nd International Conference on Computational Intelligence and Networks (CINE), Computational Intelligence and Networks (CINE), 2016 2nd International Conference on, cine. p. 112. IEEE (2016). <https://doi.org/10.1109/CINE.2016.27>, <http://search.ebscohost.com/login.aspx?direct=true&db=edsee&AN=edsee.75568>
- [15] Jeffery, D.R., Low, G.: Calibrating estimation tools for software development. Software Engineering Journal **5**(4), 215–221 (1990). <https://doi.org/10.1049/sej.1990.0024>
- [16] Heemstra, F.J.: Software Cost Estimation. Handbook of Software Engineering, Hong Kong Polytechnic University **34**(10) (1992). https://doi.org/10.1142/9789812389701_014, <https://www.st.cs.uni-saarland.de/edu/e>

- [17] Huang, X., Ho, D., Ren, J., Capretz, L.F.: Improving the CO-COMO model using a neuro-fuzzy approach. *Applied Soft Computing* **7**(1), 29–40 (jan 2007). <https://doi.org/10.1016/J.ASOC.2005.06.007>, <https://www.sciencedirect.com/science/article/pii/S1568494605000712>
- [18] Huang, S.J., Lin, C.Y., Chiu, N.H.: Fuzzy decision tree approach for embedding risk assessment information into software cost estimation model. *Journal of Information Science and Engineering* **22**(2), 297–313 (2006)
- [19] Jain, R., Sharma, V.K., Hiranwal, S.: Reduce mean magnitude relative error in software cost estimation by HOD-COCOMO algorithm. In: 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT). pp. 708–712 (2016). <https://doi.org/10.1109/ICCICCT.2016.7988044>
- [20] Chen, Z., Menzies, T., Port, D., Boehm, B.: Feature subset selection can improve software cost estimation accuracy. *ACM SIGSOFT Software Engineering Notes* **30**(4), 1 (2005). <https://doi.org/10.1145/1082983.1083171>, <http://portal.acm.org/citation.cfm?doid=1082983.1083171>
- [21] Abrahamsson, P., Moser, R., Pedrycz, W., Sillitti, A., Succi, G.: Effort Prediction in Iterative Software Development Processes – Incremental Versus Global Prediction Models. In: First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007). pp. 344–353 (2007). <https://doi.org/10.1109/ESEM.2007.16>, <https://ieeexplore.ieee.org/document/4343762>
- [22] Boetticher, G.D.: Using Machine Learning to Predict Project Effort: Empirical Case Studies in Data-Starved Domains. *Model Based Requirements Workshop* pp. 17–24 (2001). <https://doi.org/10.1.1.19.111>, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.19.111&rep=rep1&type=pdf>
- [23] Setyawati, B.R., Sahirman, S., Creese, R.C.: Neural Networks for Cost Estimation. *AACE International Transactions* p. 13.1 (jun 2002), <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=7197172&lang=de&site=ehost-live>
- [24] Finnie, G.R., Wittig, G.E.: AI tools for software development effort estimation. *Software Engineering: Education and Practice*, 1996. *Proceedings. International Conference* pp. 346–353 (1996). <https://doi.org/10.1109/SEEP.1996.534020>, <http://epublications.bond.edu.au/cgi/viewcontent.cgi?article=1006&context=infotechs>
- [25] Khalifelu, Z.A., Gharehchopogh, F.S.: Comparison and evaluation of data mining techniques with algorithmic models in

software cost estimation. *Procedia Technology* **1**, 65–71 (2012).
<https://doi.org/10.1016/j.protcy.2012.02.013>