(Thema)

Optimierung von künstlichen neuronalen Netzen zur Ausfallvorhersage mit Sensordaten

Masterarbeit

zur Erlangung des akademischen Grades “Master of Science (M.Sc.)” im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Lüdtke

Vorname: Maik

[redacted]

[redacted]

Prüfer: Prof. Dr. M. H. Breitner

Ort, den

Inhaltsverzeichnis

1. Einleitung

In unserer heutigen Zeit werden, wegen der zur Verfügung stehenden Hard- und Software, künstliche neuronale Netze immer häufiger bei der Modellerstellung zur Mustererkennungen eingesetzt. Mögliche Einsatzgebiete solcher Netzwerke sind die Bildanalyse, beispielsweise zur Gesichtserkennung [26] oder zur Detektion von Krankheiten [27], die Textanalyse [13] und viele Weitere. Aufgrund der vielfältigen Einsatzmöglichkeiten und der aktuell immer weiter steigenden praktischen Relevanz, soll sich auch diese Arbeit künstlichen neuronalen Netzen und deren Optimierung widmen. Es sollen die Daten eines Festplattendatensatzes genutzt werden, um Modelle zu entwickeln und diese mit geeigneten Methoden zu optimieren, mit dem Ziel Ausfälle dieser Festplatten im Voraus vorhersagen zu können. Dies brächte den Vorteil, dass diese Festplatten vor ihrem Ausfall repariert oder ausgetauscht werden könnten, wodurch der mögliche Verlust von Daten verhindert werden würde. Ein weiteres Ziel dieser Arbeit liegt darin, anhand des empirischen Beispieles eine mögliche Vorgehensweise zur Modellerstellung zu liefern, die auf ähnliche Anwendungsfälle übertragen werden kann. Dadurch könnte der Modellerstellungsprozess in diesen Fällen sowohl erleichtert als auch beschleunigt werden.

Um die oben aufgeführten Ziele dieser Arbeit umsetzen zu können, sollen zuerst die benötigten Grundlagen vorgestellt werden. Weil es sich bei der Kategorisierung einer Festplatte in einen Ausfall beziehungsweise Nicht-Ausfall um ein binäres Klassifikationsproblem handelt, soll zunächst die Beurteilung solcher Klassifikatoren thematisiert werden. Anschließend sollen künstliche neuronale Netze vorgestellt werden. Es soll hierbei auf den grundlegenden Aufbau, die Funktionsweise und das Training näher eingegangen werden. Des Weiteren sollen die Themen Feature Engineering und Ensemble Learning als Möglichkeiten zur Optimierung solcher Netze diskutiert werden.

Nach den Erläuterungen zu den wissenschaftlichen Grundlagen, soll die genutzte Programmiersprache *Python*, die Programme zur Erstellung von künstlichen neuronalen Netzen *Tensorflow* und *Keras*, die beide 2017 in ihren ersten stabilen Versionen veröffentlicht

worden sind, und die zur Verfügung stehenden Daten vorgestellt werden. Anschließend soll detailliert auf die empirische Arbeit eingegangen werden. Hierbei sollen alle Schritte bis zur Entwicklung des endgültigen Modells nachvollziehbar dargelegt werden. Abschließend sollen die mit diesem Model erlangten Ergebnisse diskutiert werden und es soll eine Empfehlung für den Einsatz und das weitere Vorgehen abgegeben werden.

6. Zusammenfassung

Diese Arbeit hat sich aufgrund seiner großen praktischen Relevanz primär dem Thema künstliche neuronale Netze gewidmet. Ein Ziel war es, mit Hilfe solcher Netze Modelle zu erstellen und zu optimieren, um Ausfälle von Festplatten vorhersagen zu können. Für die Erstellung der Netzwerke wurden unter anderem die äußerst aktuellen Programme Tensorflow und Keras verwendet, die, obwohl sie erst 2017 in ihrer ersten stabilen Version veröffentlicht worden sind, bereits heutzutage zu den wichtigsten Programmen in diesem Themenbereich zählen. Für die Optimierung sind verschiedenste Schritte unternommen worden, wie beispielsweise die Bearbeitung der Ausgangsdaten mit Feature Engineering, die Auswahl eines geeigneten Trainingsdatensatzes, die Auswahl von Aktivierungsfunktionen, Optimierungsalgorithmus und Verlustfunktion oder die Kombination von verschiedenen Modellen mit Hilfe von Ensemble Learning. Trotz der Umsetzung der verschiedenen Optimierungspotentiale, musste erkannt werden, dass das ursprüngliche Ziel, der Vorhersage von genauen Ausfallzeitpunkten, nicht umgesetzt werden konnte, weshalb es zu einer Umformulierung dieses Zieles gekommen ist. Ausfallzeitpunkte sollten nicht mehr genau vorhergesagt, aber so eingeordnet werden können, dass trotz einer Ungenauigkeit ein Nutzen durch die Modellverwendung generiert werden kann. Um dies bewerten zu können, wurde eine Kostenfunktion geschätzt, die Ausfälle und nicht genutzte Laufzeit, aufgrund von verfrühten Wechseln, mit Kostenfaktoren versieht. Dadurch war der Vergleich von verschiedenen Zuständen durch einen klaren Bewertungsmaßstab möglich.

Durch die Anwendung des endgültig ausgewählten Modells, konnte eine Kostenminimierung von ungefähr 20% am Validierungsdatensatz erreicht werden, welche allerdings beim Testdatensatz nicht bestätigt werden konnte. Hier kam es zu einem Anstieg der Kosten im Bereich von 4,4%, was zu der Beurteilung geführt hat, dass der Einsatz des erstellten Modells in der Realität nicht empfohlen werden kann und somit dieses gesteckte Ziel nicht umgesetzt worden ist. Es wurde angeführt, dass das erreichte Ergebnis durch die Verwendung von Stacking wahrscheinlich optimiert werden könnte, ob diese Verbesserung zu einer grundsätzlichen Änderung der Einsatzempfehlung führen würde, wird aber als eher

unwahrscheinlich eingeschätzt. Das eigentliche Problem der unzureichenden Ergebnisse wird in der lückenhaften Datenbasis gesehen, wodurch potentiell wichtige Informationen nicht oder nicht in ausreichender Qualität genutzt werden können. Aus diesem Grund wurde darauf hingewiesen, dass die vorhandenen Daten als der erfolgskritische Faktor bei der Modellerstellung eingestuft werden. Sollten Modellersteller auf die Vollständigkeit und Qualität der Daten Einfluss nehmen können, so sollte dies ihre oberste Priorität sein, weil nur dadurch sichergestellt werden kann, dass geeignete Methoden zufriedenstellende Ergebnisse liefern können. Auch wenn dies in dieser Arbeit nicht der Fall war, so wurde ein Weg der Modellerstellung dargelegt, der potentiell auf andere Anwendungsfälle abgewandelt übertragen werden kann, was dem zweiten für diese Arbeit formulierten Ziel entspricht. In welcher Qualität dieses Ziel erreicht worden ist, kann an dieser Stelle nicht beurteilt werden, weshalb dies dem jeweiligen Leser überlassen wird.