University of Bergen
Department of Informatics

GNN

Author: Bendik Akselsen Solevåg
Supervisors: Ramin Hasibi and Tom Michoel

UNIVERSITETET I BERGEN
Det matematisk-naturvitenskapelige fakultet

February, 2023

Abstract

Abstract

Acknowledgements

Acknowledgements

Bendik Solevåg
Wednesday 1st February, 2023

Contents

1 Introduction **1**

2 Background **2**

2.1 Machine Learning 2

2.1.1 What is machine learning? 2

2.1.2 Linear Regression 3

2.1.3 Multi Layer Perceptron 4

2.1.4 Neural Networks 4

2.1.5 Backpropagation 5

2.1.6 Supervised and unsupervised learning 5

2.2 Graphs 5

2.2.1 Nodes and Edges 5

2.2.2 Directed Acyclic Graphs 5

2.3 Graph Neural Networks 5

2.3.1 The Convolution Layer 5

2.3.2 Permutation invariance and equivariance 5

2.3.3 Attention and Message Passing 5

2.4 Gated Recurrent Units 5

2.4.1 Backpropagation Through Time 5

3 Project methodology and setup **6**

3.1 The Geauvadis Dataset 6

Bibliography **7**

List of Figures

2.1 Linear regression illustrated 4

List of Tables

Listings

Chapter 1

Introduction

Chapter 2

Background

This chapter aims to explain the basic concepts that are the most important to the work in this thesis.

2.1 Machine Learning

The following sections will define the term 'Machine Learning'. It will describe some various ways that machine learning can be applied.

2.1.1 What is machine learning?

As a term, 'Machine Learning' is a subcategory if the umbrella term 'Artificial Intelligence'. It is proposed as an alternative to traditional algorithms. Machine Learning takes a set of input (training) data, and attempts to reason about some quality of the input, without the author of the program explicitly telling the program what quality of the input data we are interested in reasoning about. Instead, the author provides the Machine Learning model with their optimal target for the given input, and the model must attempt to generalize over, and design its own algorithm to fit the target.

This approach has become useful in problems where discovering the target based on the model input becomes computationally intractible, or when the target cannot be determined as a direct consequence of the input. An example of such a problem is

sentiment analysis of text input. Given the sentence 'The nice boy made fun of the kind girl', it is difficult to design rules which can capture the sentiment of the sentence. On one hand, the words 'nice', 'fun', and 'kind' indicate this sentence may have a positive sentiment. In reality, our ability to reason tells us that this is not the case, and the sentence is of a negative nature.

Machine Learning is deeply rooted in Bayesian Statistics. (Noen gode setninger om bayes. Gjerne bayes teorem? Ikke mer enn et avsnitt. Dette er en AI-oppgave, ikke en statistikk-oppgave.)

2.1.2 Linear Regression

As a first study in the implementation of machine learning, we look at linear regression. To explan linear regression, we view an example in a two dimensional, euclidean space. We are given a set of input data values X . Each input value $x_i \in X$ has a corresponding output value $y_i \in Y$. We can plot each input value, and its corresponding output value as coordinates.

To describe a function that best maps the given inputs with their related outputs, we introduce two variables w_0 and b_0. We define a function to map our input values to their corresponding output.

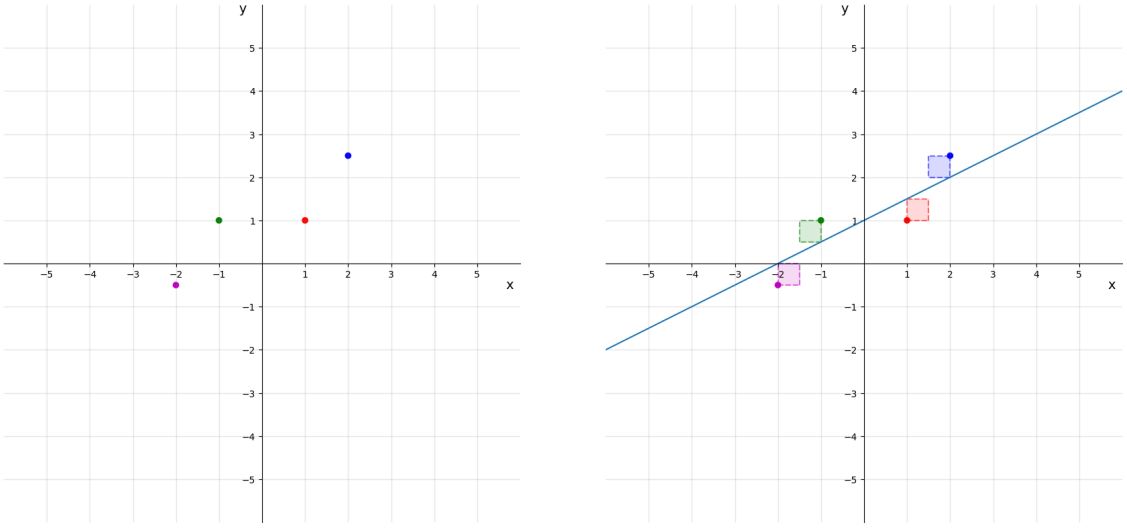
$$f(x) = w_0 x + b \tag{2.1}$$

We also introduce some measure of how well our function (2.1) maps our input to our output. As we tune our function variables w_0 and b_0 to better map the input to its corresponding output, the measure will decrease. There exists values for our variables w_0 and b_0 such that the measure defined in (2.2) reaches its lowest possible value. In this state, the function optimally maps the input data to with its corresponding output, given our function.

$$\mathbb{L} = \frac{1}{|X|} \sum_{i=0}^{|X|} \big(y_i - f(x_i)\big)^2 \tag{2.2}$$

Of course, the function defined in (2.1) is arbitrary. We can add or remove existing or additional function variables as we wish, if we see that changing the function structure improves at mapping the input to the output.

Referanse til taylor series expansion? Finn i Calculus bok



(a) Input x values plotted against their corresponding y values

(b) label 2

Figure 2.1: Linear regression illustrated

Loss

In the previous section, a measure of how 'wrong' a mapping of input points to their corresponding output points was defined. In the litterature, this measure is referred to as the loss function. The loss function defined in (2.2) is named 'Mean Squared Error Loss', due to the fact that it is defined by the mean of the squared difference between our function's predictions and the true values.

It is possible to define different loss functions. Different loss functions will interpret the performance of a functions mapping of the given input data to its corresponding output in different ways. If we

$$\mathbb{L} = \frac{1}{|X|} \sum_{i=0}^{|X|} \sqrt{\big(y_i - f(x_i)\big)^2} \tag{2.2}$$

2.1.3 Multi Layer Perceptron

2.1.4 Neural Networks

underfitting and overfitting?

Linear Layers

Activation functions

2.1.5 Backpropagation

2.1.6 Supervised and unsupervised learning

2.2 Graphs

2.2.1 Nodes and Edges

2.2.2 Directed Acyclic Graphs

2.3 Graph Neural Networks

2.3.1 The Convolution Layer

2.3.2 Permutation invariance and equivariance

2.3.3 Attention and Message Passing

2.4 Gated Recurrent Units

2.4.1 Backpropagation Through Time

Chapter 3

Project methodology and setup

3.1 The Geauvadis Dataset

Bibliography