

Diagnostisering av cancer med faltningsnätverk

William Nylund 41686

Kandidatavhandling i datateknik

Handledare: Johan Lilius

Fakulteten för naturvetenskaper och teknik

Åbo Akademi

2021

Referat och nyckelord

TODO

Innehållsförteckning

1	Inledning	3
2	Maskininlärning	4
2.1	Inlärningsmetoder	4
2.1.1	Övervakad inlärning	4
2.1.2	Oövervakad inlärning	4
2.1.3	Indelning av data	5
2.2	Generalisering och överanpassning	5
3	Introduktion till artificiella neuronät	7
3.1	Neuronät	7
3.2	Artificiella neuronät	8
3.2.1	Den artificiella neuronerna	8
3.2.2	Det artificiella neuronätet	9
3.3	Bildigenkänning	9
3.4	Faltningsnätverk	10
3.5	Bakåtpropagering	10
4	Tillämpning av faltningsnätverk vid cancerdiagnostisering	11
4.1	Klassificering av hudsjukdomar	11
4.2	Identifiering av bröstcancer	13
4.3	Riskbedömning av lungcancer	14
5	Diskussion	15
6	Sammanfattning	16

1. Inledning

TODO

2. Maskininlärning

Artificiella neuronnät är en underkategori till maskininlärning, som i sin tur är en underkategori till Artificiell intelligens (AI). En vanlig definition av AI är ett system som tänker och agerar rationellt, eller system som tänker och agerar likt människor [15]. Ett sätt för ett system att uppnå detta är genom maskininlärning. Maskininlärning är ett område inom datavetenskap som försöker bygga en *modell* som simulerar ett fenomen. Man lär upp modellen baserat på befintlig träningsdata [2]. Denna data består av *datapunkter* som är ett element i mängden, t.ex. en person. Modellen finner mönster i data och kan sedan göra förutsägelser på främmande data. Det finns olika typer av inlärningsmetoder som väljs beroende på typen av problem man löser samt vilken typ av data man har tillgänglig.

2.1 Inlärningsmetoder

2.1.1 Övervakad inlärning

Övervakad inlärning (eng. supervised learning) är den mest använda inlärningsmetoden [9]. I övervakad inlärning tränar man modellen med data som är par av indata och utdata. Modellen kan till exempel få personer som indata och ge antalet armhävningar personen kan göra som utdata. Modellen tränas först på träningsdata, som har *etiketter* (eng. labels), d.v.s. det förväntade utdata man vill få ut av modellen, t.ex. antalet armhävningar personen kan göra. Etiketter fungerar som facit för modellen. Indata har *egenskaper* (eng. features). En egenskap är något som beskriver datapunkterna, Detta kan till exempel vara personens längd och vikt. Målet är att modellen kan räkna ut hur en persons längd och vikt relaterar till hur många armhävningar den kan göra och sedan kunna förutspå antalet armhävningar en person kan utföra med endast dess längd och vikt till förfogande.

2.1.2 Oövervakad inlärning

I *oövervakad inlärning* (eng. unsupervised learning) har man däremot inte definierat några etiketter (förväntad utdata). Istället visar man data med egenskaper åt sys-

temet och låter det extrahera information på egen hand [9]. På detta sätt kan man t.ex. hitta nya sätt att kategorisera data i s.k. *kluster*. Ett kluster är en grupp av datapunkter som har liknande egenskaper.

2.1.3 Indelning av data

När man utvecklar maskininlärnings system delar man in den tillgängliga data i olika mängder som var och en uppfyller en egen funktion. Man delar oftast in data i tre mängder:

- Träningsdata
- Valideringsdata
- Testdata

Av dessa mängder utgör oftast träningsdata den största [2]. Träningsdata används när man tränar modellen. Efter träningsperioden provar man modellen på valideringsdata. Beroende på det resultatet går man tillbaka och modifierar inlärningen. Testdata användes slutligen för att bedöma modellen innan man levererar produkten [2]. En orsak till uppdelningen av data är för att kunna testa hur väl modellen *generaliserar* från en datamängd till en annan.

2.2 Generalisering och överanpassning

I övervakad inlärning vill man utveckla en modell baserat på träningsdata som sedan kan göra korrekta antaganden på okänd data [9]. Om modellen kan göra detta säger man att den generaliserar väl från en datamängd till en annan, vilket är en egenskap man eftersträvar hos en modell. Det visar sig att bygga en modell som generaliserar bra blir en balansgång mellan *underanpassning* (eng. underfitting) och *överanpassning* (eng. overfitting). En modell är underpassad om den gör många misstag på träningsdata [2]. Orsaken till underanpassning kan vara att modellen är för enkel eller att egenskaperna av datapunkterna inte är tillräckligt informativa [2]. Underanpassning kan inträffa om man till exempel försökte lära ut en modell att förutspå armhävningar, men de enda egenskaperna tillgänglig var längd. Underanpassning illustreras till vänster i figur 2.1. Där har man försökt anpassa modellen med en linjär kurva medan den i själva verket skulle ha behövt en mera komplex modell. På andra sidan spektrumet finns överanpassning. En modell blir istället överanpassad när modellen är för komplex för den data man har till förfogande eller om man har för få datapunkter [2]. En överanpassad modell är så specialiserad på träningsdata att den inte generaliserar väl till ny data. I figur 2.1. är en

överanpassad modell illustrerad till höger. Den använder sig av en femtondegradsekvation medan den i själva verket skulle ha generaliserat bäst till ny data med en andragsradsekvation som är illustrerad i mitten av figur 2.1.

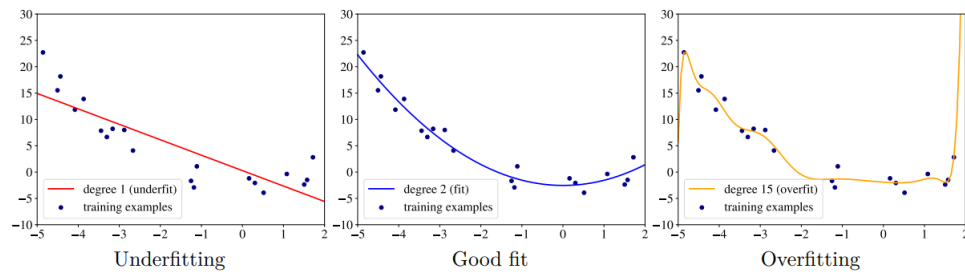


Figure 2.1: Exempel på underanpassning och överanpassning [2]

3. Introduktion till artificiella neuronnät

3.1 Neuronnät

Efter en grundläggande genomgång av maskininlärning kommer denna avhandling först att introducera *biologiska neuronnät* och sedan *artificiella neuronnät*. För att kunna utveckla maskiner som tänker och agerar likt eller bättre än människor, började man undersöka människans hjärna i hopp om att kunna imitera den. Warren S. McCulloch och Walter Pitts presenterade 1943 en matematisk modell av hur hjärnans neuronnät fungerar [8]. Hjärnan består av 10^{11} (100 miljarder) *neuroner* eller *nervceller* [6]. Dessa neuroner kommunicerar med varandra genom kemiska signaler som kallas *synapser*. Neuronen tar emot signaler från andra neuroner genom *dendriterna*. Dendriterna är av olika grad *hämmande* eller *exciterande* och beroende på storleken av dessa egenskaper avfyra neuronerna en utsignal genom *axonen*. Denna utsignal blir i sin tur insignal till andra neuroner. Neuronen illustreras i figur 3.1. Nervsystemet har haft stort inflytande i utvecklingen av artificiella neuronnät och likheterna kommer att bli tydliga i följande kapitel.

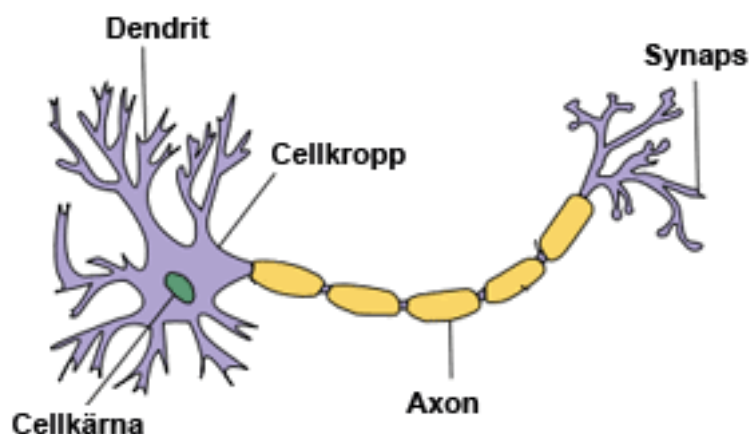


Figure 3.1: Neuronens uppbyggnad [4]

3.2 Artificiella neuronnät

3.2.1 Den artificiella neuronen

Likt vårt nervsystemet har det artificiella neuronnätet *noder* som motsvarar neuroner. Noderna är sammanlänkade som neuronens dendritter och axoner med flera insignaler och en utsignal. Kopplingarna är likt synapserna av olika grad exciterande eller hämmande, som beskrivs av en vikt w som oftast är ett positivt eller negativt decimaltal [6]. Om noden får input av n stycken noder så är x_1, x_2, \dots, x_n nodens insignaler och w_1, w_2, \dots, w_n är kopplingarnas vikter. Då beräknas nodens totala summa av

$$summan = w_1x_1 + w_2x_2 + \dots + w_nx_n.$$

Denna summa ges sedan som input till en aktiveringsfunktion som beräknar outputen. En grundläggande aktiveringsfunktion är den binära stegfunktionen. Den jämför summan med ett gränsvärde θ och utsignalen produceras antingen som ett eller noll med hjälp av stegfunktionen [13]

$$f(x) = \begin{cases} 1, & \text{då } summan > \theta \\ 0, & \text{då } summan \leq \theta \end{cases}.$$

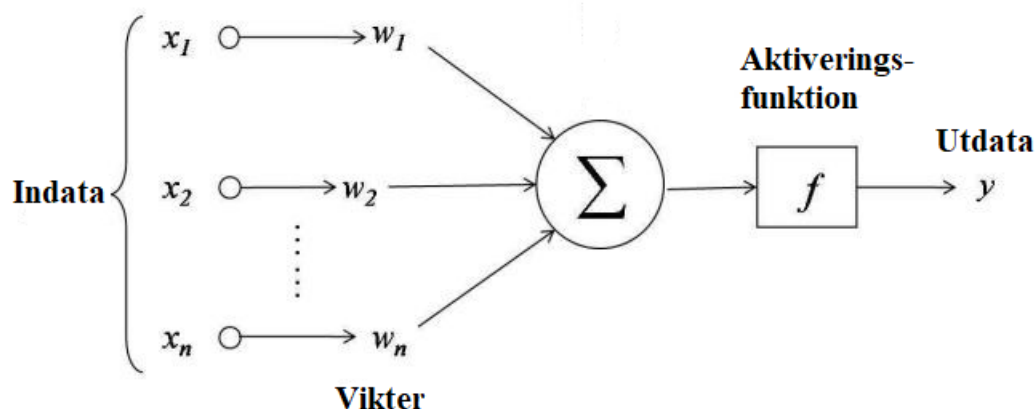


Figure 3.2: Den artificiella nodens uppbyggnad [12]

3.2.2 Det artificiella neuronnätet

Förra kapitel gick vidare igenom den artificiella neuronens uppbyggnad. Nästa steg är att kombinera dem för att bilda ett neuronnät. Arkitekturen i ett neuronnät består av lager av noder. Antalet lager varierar men det finns alltid ett indatalager (eng. input layer) och ett utdatalager (eng. output layer). Mellan dessa finns gömda lager vars antal varierar mellan nätverk. Nätverk med flera än ett gömt lager kallas djupa neuronnät [11]. Ett nätverk kan inte innehålla cyklar utan information rör sig från indatalagret till utdatalagret genom varje lager [13]. Ett nätverk är illustrerat i figur 3.3. Kanterna mellan noderna har vikter och beräkningar sker i noderna enligt kapitel 3.2.1.

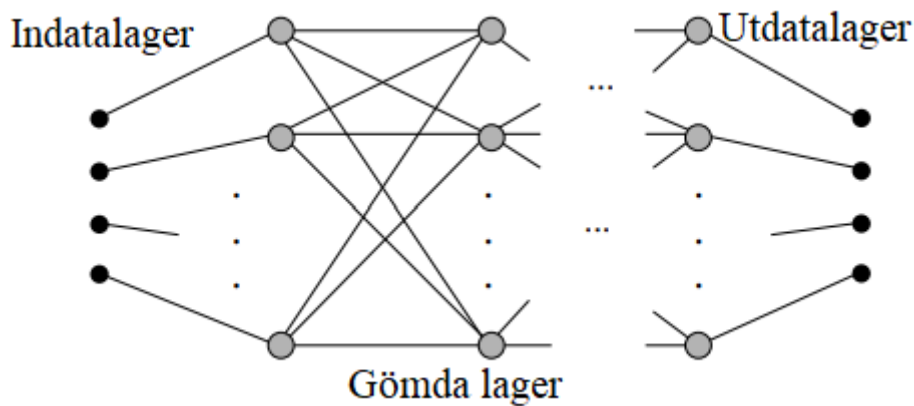


Figure 3.3: Det artificiella neuronnätets arkitektur [13]

3.3 Bildigenkänning

Ett sätt att diagnostisera patienter med hjälp av artificiella neuronnät är genom bildigenkänning. Målet är att man ska kunna mata in en bild av t.ex. ett födelsemärke som indata och att nätverket ska kunna säga oss om födelsemärket är cancerartad eller inte. Ett nätverk analyserar en bild genom att varje pixel korresponderar till en nod i indatalagret [11]. I en svartvit bild representeras en pixel av ett heltal mellan 0 och 255, där 0 är helsvart och 255 helvit. I en bild med färg representeras pixeln enligt *RGB-modellen*, d.v.s. tre heltal som representerar hur mycket av rött, grönt eller blått pixeln innehåller. Istället för att undersöka varje pixels förhållande till resten av pixlarna använder man sig av *faltningsnätverk*.

3.4 Faltningsnätverk

Faltningsnätverk (eng. convolutional neural network) är en typ av neuronnät som ofta används i bildigenkänning för att hitta egenskaper i bilder t.ex. trafikljus eller bokstäver. Hittills har lagren i nätverket representerats som en vertikal linje av noder. I ett faltningsnätverk underlättar det att istället representera lagren som en kvadrat där varje pixel har en nod. I följande lager kommer varje nod vara kopplat till en region, t.ex 5x5 noder, av indatalagret [11]. Detta görs genom att välja ett s.k. *filter* eller *kärna* (eng. kernel). Bilden skannas med hjälp av filtret för att hitta önskade egenskaper i bilden. Detta illustreras i figur 3.4.

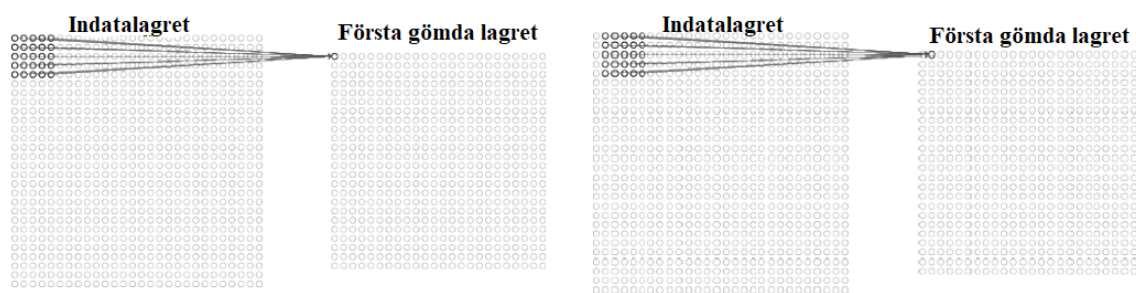


Figure 3.4: Filtret går igenom indatalagret och skapar det första gömda lagret [11]

3.5 Bakåtpropagering

Bakåtpropagering (eng. backpropagation) är en övervakad inlärningsmetod som ofta används inom neuronnät [11]. Namnet kommer ifrån att man går genom nätverket baklänges och modifierar vikterna tills man får det önskade resultatet. Metoden börjar med att man sätter vikterna slumpmässigt [6]. Sedan beräknas felet (eng. error) utifrån etiketten. Man går då tillbaka och justerar vikterna för att närma sig det korrekta värdet. Detta är en iterativ process som utförs tills felet är tillräckligt litet.

4. Tillämpning av faltningsnätverk vid cancerdiagnostisering

4.1 Klassificering av hudsjukdomar

Andre Estava et al. [5] utvecklade ett neuronät som kan klassificera olika hudsjukdomar från bilder. Hudcancer är den tredje vanligaste cancersjukdomen [3] och processen att identifiera hudcancer är lång och kan ta flera läkarbesök till dermatologer. Diagnostiseringen sker primärt visuellt [5], vilket gör det en utmärkt kandidat för ett använda faltningsnätverk.

Andre Estava et al. använde GoogleNet Inception v3 CNN arkitekturen [16] som var tränat på förhand på databasen Imagenet [14], 1.28 miljoner bilder som tillhör 1000 olika kategorier. Med hjälp av Inception v3 tränade de ett nätverk med *överföringsinlärning* (eng. transfer learning) på ny data. Överföringsinlärning används ofta när man utvecklar faltningsnätverk. Det går ut på att man använder ett nätverk som redan är tränat på ett stort dataset, ofta Imagenet [14], och använder sig av dess tidigare lager som fokuserar på de grundläggande egenskaperna i bilden [7]. De sista lagren byts ut för att specificera sig på de egenskaper man vill identifiera. En översikt av nätverket illustreras i figur 4.1.

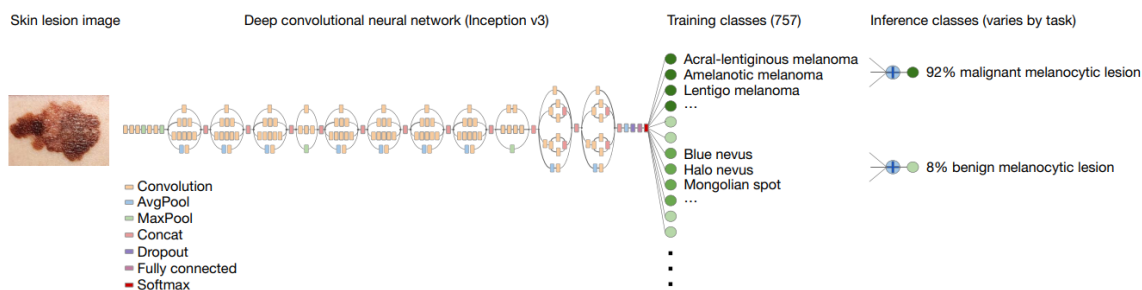


Figure 4.1: Faltningsnätverkets layout [5]

För att träna de sista lagren i nätverket använde Andre Estava et al. 129,450 bilder som dermatologer hade satt etiketter på. Dessa delades in så att 127,463 var träningsdata och valideringsdata medan 1,942 bilder användes som testdata

[5]. Bilderna organiserades enligt en trädstruktur bestående av 2,032 hudsjukdomar där löven är individuella sjukdomar och deras förälder är samlingsnamn [5]. En delmängd av trädet illustreras i Figur 4.2.



Figure 4.2: Träd av hudsjukdomar [5]

Man jämförde faltningsnätverkets och två dermatologers utförande i två olika uppgifter. Den första var att klassificera första nivån i trädet, d.v.s. om hudsjukdomen är godartad (eng. benign), elakartad (eng. malignant) eller icke-neoplasiska (eng. non-neoplastic). Detta gjordes genom att summera sannolikheterna för dessa noders barn [5]. Här hade nätverket i medeltal 72.1% rätt med en standard avvikelse på 0.9% medan dermatologernas medeltal var 65.56% och 66.0% [5]. I den andra uppgiften skulle de klassificera andra nivån i trädet. Här fick nätverket i medeltal 55.4% rätt med standard avvikelsen 1.7%. Även i denna uppgift presterade dermatologerna lite sämre än nätverket med 53.3% och 55.0% rätt[5].

Andre Estava et al. bevisade att nätverket kunde prestera på samma nivå som certifierade dermatolger. Ett användningsområde för nätverket är i en telefonapplikationen [5]. Patienten kunde göra en förundersökning på egen hand med sin telefon och få reda på om hen behöver besöka en läkare, och i så fall, boka en tid till rätt specialist. Detta skulle spara båda läkarens och patientens tid. Applikationen kunde också användas i länder där det finns stor brist på dermatologer.

4.2 Identifiering av bröstcancer

The International Symposium on Biomedical Imaging ordnade en tävling i bildigenkänning 2016. Första delen av tävlingen var att identifiera om huruvida en bild innehöll bröstcancer, medan den andra var att lokalisera var i bilden canceren är positionerad. Dayong Wang et al. [17] vann båda med deras faltningsnätverk.

Den data som var tillgänglig under tävlingen var 400 bilder som var skannade med ett mikroskop för att skapa en enda digital bild i hög upplösning. Bilderna delades in i 270 för träning och 130 för testning. Bilderna hade etiketter som beskrev huruvida bilden innehöll cancer och i så fall regionen där cancer fanns. Etiketterna var placerade av en patolog [17]. Under träningen valde man slumpmässigt miljontals regioner bestående av 256×256 pixlar av bilden och som fick etiketter med ett eller noll beroende på om regionen var cancerartad eller inte [17]. Sedan tränade man modellen på dessa regioner för att känna igen cancerartade regioner. En värmekarta illustrerades för att visa sannolikheten att regioner är cancerartade Dayong Wang et al. Testade olika nätverks nätverksarkitekturer men GoogleNet Inception [16] presterade bäst. En illustration av träningsperioden är illustrerad i figur 4.3.

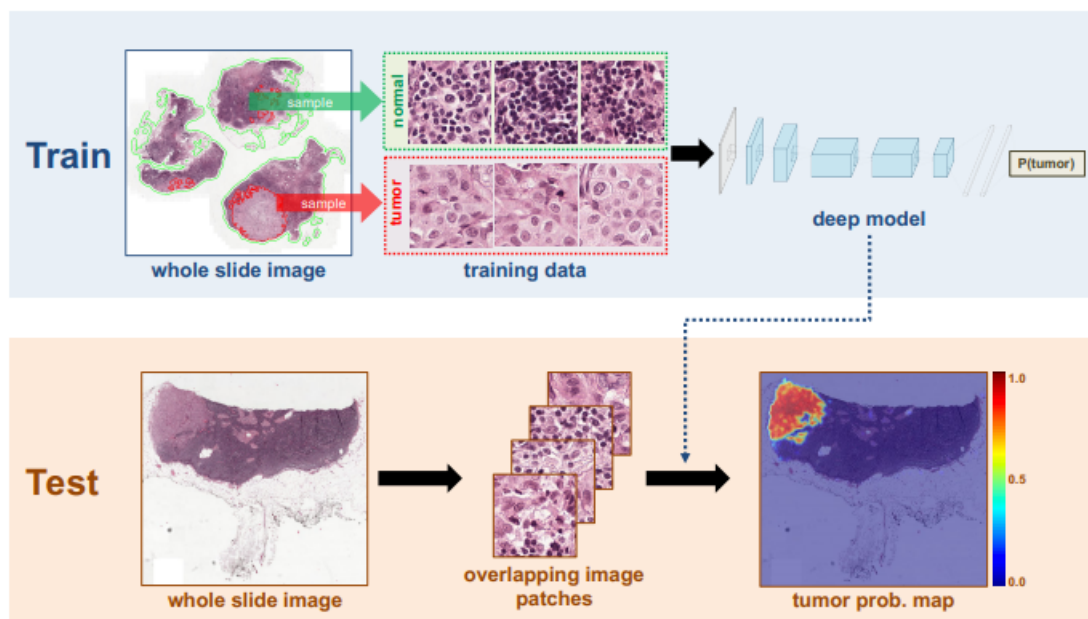


Figure 4.3: training [17]

Resultaten i tävlingen presenterades i form av *AUC* (eng. area under the curve) [17]. *AUC* är ett av de vanligaste mått av prestanda i maskininlärning [10]. *AUC* är ett värde mellan noll och ett som berättar hur bra modellen kan berätta om huruvida bilden innehåller cancer eller inte. Ett *AUC* värde som är ett är önskvärt, d.v.s. modellen kan alltid identifiera bilder som innehåller cancer. Noll betyder att

modellen alltid förutspår motsatsen. Medan ett halvt är den värsta situationen då modellen inte alls kan skilja på bilderna [10].

Dayong Wang et al. uppnådde ett AUC värde på 0.9250 medan en patolog som undersökte samma bilder fick 0.9664 [17]. Det som dock var intressant i denna studie var att felen som faltningsnätverket och patologen gjorde skiljde sig avsevärt, vilket betyder att om patologen skulle ta hjälp av nätverket i sitt arbete skulle de komplettera varandra och tillsammans uppnå ett AUC värde på 0.9948 [17].

4.3 Riskbedömning av lungcancer

TODO: Ardila, D., Kiraly, A.P., Bharadwaj, S. et al. [1]

5. Diskussion

TODO

6. Sammanfattning

Referenser

- [1] Kiraly A.P. Bharadwaj S. et al. Ardila, D. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. page 954–961, 2019.
- [2] Andriy Burkov. *The Hundred-Page Machine learning Book*. 2019.
- [3] Cancerfonden. De vanligaste cancersjukdomarna. 2018.
- [4] Svenskt Demenscentrum. Vad händer i hjärnan? 2020.
- [5] Kuprel B. Novoa R. et al. Esteva, A. Dermatologist-level classification of skin cancer with deep neural networks. page 115–118, 2017.
- [6] Kevin Gurney. *An Introduction to Neural Networks*. CRC Press, 2003.
- [7] Will Koehrsen. Transfer learning with convolutional neural networks in pytorch. 2018.
- [8] Pitts W McCulloch, W.S. A logical calculus of the ideas immanent in nervous activity. bulletin of mathematical biophysics 5. page 115–133, 1943.
- [9] A.C. Müller and S.Guid. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Inc, 2016.
- [10] Sarang Narkhede. Understanding auc - roc curve. 2018.
- [11] Michael Nielsen. *Neural Networks and Deep Learning*. Springer Science Business Media, 2013.
- [12] Prateek. Statistics is freaking hard: Wtf is activation function. 2018.
- [13] Raul Rojas. *Neural Networks: A Systematic Introduction*. Springer Science Business Media, 2013.
- [14] O. et al Russakovsky. Imagenet large scale visual recognition challenge. page 211–252, 2015.

- [15] Stuart J. Russell and Peter Norvig. *Artificial Intelligence, a Modern Approach*. Prentice Hall, 1994.
- [16] Vanhoucke V. Iofe S. Shlens J. Wojna Z Szegedy, C. Rethinking the inception architecture for computer vision. 2015.
- [17] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck. Deep learning for identifying metastatic breast cancer. 2016.