**Reflectie-verslag: Implementatie van een Denoising Diffusion Probalistic Models**

**Studenten:**

Julia Boschman

Daan Eising

**Studentnummers:**

22074845

22162038

**Datum:**

23 maart 2025

1. **Inleiding**

Voor het vak Gen AI hebben wij, Julia Boschman en Daan Eising, ervoor gekozen om een Denoising Diffusion Probalistic Model te implementeren aan de hand van een dataset die Baroque stijl schilderijen beelden bevat, gebaseerd op het paper “Denoising Diffusion Probabilistic Models”. Het doel was om het model volledig ‘from scratch” te programmeren. Hierbij mochten wij geen packages in Python gebruiken om het model uit te voeren. In dit verslag benoemen wij de technische uitdagingen en leermomenten die wij gedurende het maken van het portfolio zijn tegengekomen.

1. **Problemen en oplossingen**

Gedurende het make van het portfolio hebben wij last gehad van meerdere problemen. Hieronder benoemen wij deze problemen en de uiteindelijke oplossingen die wij hebben ingevoerd om zo goed mogelijk deze problemen te verhelpen.

**Starten**

Het eerste probleem waar wij als duo tegen aan liepen was het starten van het portfolio. Na de eerste les van het vak hebben wij samen een planning opgesteld. Wij deden dit om ervoor te zorgen dat wij het portfolio vroegtijdig zouden afronden. Echter de vooropgesteld planning bleek niet realistisch en hierdoor lastig om aan te houden. De voornaamste reden was de bijbehorende paper, zo was voor ons het lezen en het begrijpen van de paper een erg grote drempel. Hierbij kwam ook een andere bij zaak, de nieuwe stof en de grote hoeveelheid huiswerk dat los stond van het portfolio. Wij waren deze belasting niet gewend en hadden tijd nodig om ons hierop aan te passen. De combinatie hiervan heeft ervoor gezorgd dat wij het portfolio uit het zicht zijn verloren.

**Oplossing voor het opstart probleem.**

De oplossing voor het probleem is tot stand gekomen in lesweek 4. Na wat aanpassingen aan het vak hadden wij meer ademruimte en kwamen wij tot de conclusie dat wij nu stappen moesten maken in het maken van het portfolio. Zo hebben wij als duo afgesproken om actiever onderling te communiceren, duidelijke afspraken te maken. Een van de afspraken was dat wij als duo iedere school dag na de contacturen samen op een rustige locatie zouden werken aan het portfolio. Dit heeft ervoor gezorgd dat wij daadwerkelijke zijn begonnen en grote stappen de juiste richting op hebben genomen.

**Complexiteit paper**

Toen wij uiteindelijk in lesweek 4 met meer focus aan het portfolio begonnen lukte het ons beide niet om de paper eigen te maken en daadwerkelijk te begrijpen. Zo hadden wij in week 1 gekozen voor de porfolio opdracht: Normalising flow model op basis van “Variational Inference with Normalizing Flows”. Na meerdere pogingen is het ons ook niet gelukt om de code “from scratch” te implementeren.

**Oplossing complexiteit paper**

Door deze tegenslag van het implementeren en begrijpen was onze motivatie flink gedaald. Uiteindelijk hebben wij besloten dat het “from scratch” implementeren op basis van deze papier niet ging lukken. Vervolgens hebben wij de overige papers bekeken om een ervan te selecteren voor onze tweede poging. Hieruit sprak de paper “Denoising Diffusion Probabilistic Models” ons erg aan. Zo waren wij in dezelfde periode als het portfolio ons ook aan het voorbereiden op de paper presentatie over causal fusion. Deze paper had veel raakvlakken met de paper over DDPM. Hierdoor begrepen wij de paper beter, hadden wij meer ideeën voor mogelijke datasets en keerde onze motivatie terug. Eindstand hebben wij na deze keuze harder en met meer enthousiasme aan het portfolio gewerkt.

**Wiskundige concepten**

Na het veranderen van paper begrepen wij het concept beter. Echter bevatten de paper geavanceerde wiskunde wat het ons lastig maakte om over te zetten naar code. Vooral de formules en achterliggende concepten zorgden ervoor dat we veel tijd moesten besteden aan het begrijpen ervan.

**Oplossing moeilijke wiskundige concepten**

Om alsnog de wiskunde goed te kunnen omzetten hebben wij externe bronnen geraadpleegd om de concepten en formules in kleine delen op te splitsen en zo beter te begrijpen. De grootse hulp hierbij was chatgpt. Zo was het met deze tool mogelijk om alles wat we niet begrepen steeds simpeler uit te laten leggen, totdat we het begrepen. Zo konden werd het alsnog mogelijk om de informatie uit de paper eigen te maken.

**Trainingstijd**

Toen wij eindelijk het model geimplementeerd hadden kwam nog een probleem aan het licht. Door het “from scratch” implementeren van het model hadden wij niet de gebruikelijke optimalisatie technieken toegepast die normaal standaard zitten inbegrepen in keras en pytorch. Zo konden wij nu ook niet het model laten draaien via een GPU maar ging nu alles via numpy en werd alleen de CPU benut. Hierdoor was de trainingstijd van het model aanzienlijk langer dan verwacht. Zo duurde het trainen van een epoch ongeveer 20 minuten. Met deze snelheid zou het trainen van het model te veel tijd in beslag nemen.

**Oplossing trainingstijd**

Om alsnog het model goed te kunnen trainen hebben wij het model geoptimaliseert door vectoriasatie te te passen. Verder hebben wij ook de afbeelding groote van 128x128 verkleind naar 32x32. Uiteindelijk heeft dit ervoor gezorgd dat de trainingstijd per epochj uiteindelijk terug gebracht kon worden naar ongeveer 1 minuut. Nu was het mogelijk om daadwerkelijk het model te trainen en ook efficiënter verschillende parameters te testen en aan te passen.

**Gegenereerde samples**

Na het uitvoeren van meerdere test’s en verschillende trainingsloops bleef het model afbeeldingen generen waarbij alleen de toegevoegde ruis te zien was. Daarbij bleef de bijbehorende ruis erg hoog, tijdens onze beste trainingsloop was het laagste dat we gezien hebben 0.9. Verder bleven de gemiddelde gradients erg. Dit gaf dus aan dat het onvoldoende leerde en dat er wat miss gaat bij de reverse process.

**Oplossing gegenereerde samples**

We hebben uitgebreid geexperimenteerd en geprobeerd meerdere verschillende vormen van reverse process toe te passen. Echter is het ons niet gelukt om de loss te verkleinen en het reverse process te verbeteren. Eindstand is het ons dus niet gelukt om dit probleem op te lossen. In de toekomst zullen wij een soortgelijk probleem proberen te voorkomen of op te lossen door eerder te starten en meer hulp te vragen aan de vak docent.

**Onderschatting complexiteit en benodigde tijd**

Hoewel de opdracht interresant en erg leerzaam was bleek uiteindelijk de complexiteit en de hierbij benodigde tijd groter dan verwacht. Het voor het eerst volledig “from scratch” een model uit een paper bouwen en trainen op nieuwe data was erg uitdagend en tijdrovend.

**Oplossing**

Voor de komende portfolio’s en in de toekomst soortgelijke projecten zullen wij meer tijdreserveren voor de opdracht en hierbij een realistischere planning maken. Zo willen wij voorkomen dat wij een niet volledig geimplementeerd model opleveren.

1. **Nieuwe bevindingen**
2. **Conclusie**

**Tips voor toekomstige implementaties**

1. **Bronnen**

De Vries, J. (2024, 26 januari). *Reflectie-verslag: Implementatie van een score-based model* [Reflectieverslag]. De Haagse Hogeschool. <https://brightspace.hhs.nl/content/enforced/121890-ADSAI-GENAI-24_2024_SPRING_2/voorbeeldreflectie.pdf?ou=121890>

Of

Reflecti-verslag voorbeeld afkomstig van brightspace Gen AI.