Wrap Up

ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

SCHMIDT FUTURES

Takeaways?

- What takeaways do you have from today?

Best Practices: Model Choices

- Be intentional about the choice between Open and Closed models and the tradeoffs
  - Closed models are easy to use, but we cannot know exactly what is taking place behind the API
  - Open source models are easier to control, though we still may not know exactly how they were trained
- If you are using models trained by other people:
  - Understand as much as you can about how their LLM was trained, what sort of data was used, and how it was finetuned
- Consider using a small language model – you may not need a large model for your task
- Consider using a quantised model, for energy efficiency
- Consider using non-AI techniques
- Check the licenses of data, models & code that you use

Best Practices: Experimentation

- Build a baseline model using the simplest techniques you can
- Ensure reproducibility of your experiments where possible
  - Use temperature parameter to control randomness
  - Use an open source model
- Use the University's computing facilities – HPC & Dawn GPUs – when you need computing power
- Use Data Sheets and Model Cards for documentation

Best Practices: Evaluation

- Carefully specify the task you want your model to perform
- Create a test dataset which you can use to evaluate the performance of any AI model
  - Ensure that your test dataset is separate from any training data
- Consider data leakage and how your test data might already be in the model's training data
- Understand the limitations and common issues with LLMs, e.g.
  - Bias
  - Hallucination
- Understand the metrics you're computing and what they can tell you

Other Accelerate Material

Get started

large-language-models

Get an introduction to how large language models work, and how to get up and running quickly.

Jupyter Notebook ☆ 12 7

packaging-publishing

Learn how to package and publish your scientific research software using GitHub Actions.

Jupyter Notebook ☆ 10 4

data-school-Autumn23

An introduction on how to build effective data pipelines for machine learning projects.

Jupyter Notebook ☆ 4

diffusion-models

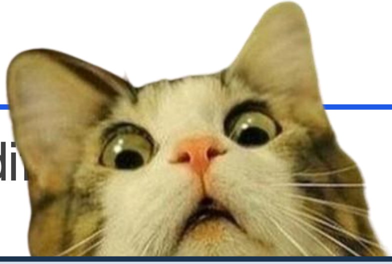
An introduction to diffusion models and how to apply them in a practical way

Jupyter Notebook ☆ 5 1

ALL REPOS

Diffusion Models

Here is the algorithm for di



Algorithm 1 Training

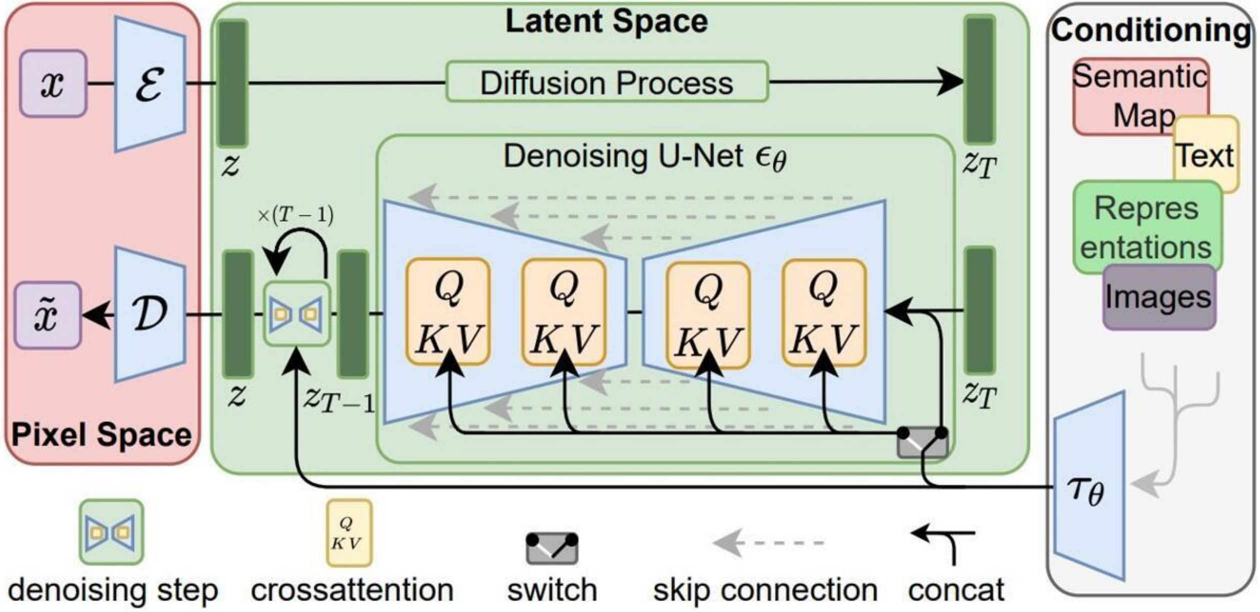
repeat
  $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
  $t \sim \mathrm{Uniform}(\{1, \ldots, T\})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  Take gradient descent step on
  $$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|^2$$
until converged

Algorithm 2 Sampling

$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, \ldots, 1$ **do**
  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
  $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t \mathbf{z}$
end for
return $\mathbf{x}_0$

Diffusion Models

Perhaps this makes more sense… maybe not…



ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

SCHMIDT FUTURES

Upcoming events and workshops

Teaching workshops

Monday 14 October	Hands on LLM workshop
Wednesday 6 November	An Introduction to Diffusion Models in Generative AI
Monday 18 November	Publishing and Packaging Python Code for Research
Monday 2 December	An Introduction to Docker
Wednesday 4 December	Hands on AI workshop

Sign up through our events page:

bitly

AI Cafés

Wednesday 23 October – West Hub

Monday 11 November – St Edmund's College

Get in touch for further information and support:

accelerate-mle@cst.cam.ac.uk

AP SCI ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

SCHMIDT FUTURES