Task 3.2.1

Imagine using a random prompt like "a cat riding a bicycle on Mars" and seeing an AI generate an image that matches your description perfectly. This is made possible by using advanced models like DALL-E, which use various machine learning techniques, including the diffusion processes. For this task, explain what basic diffusion is, how it works, and why it is used in generating such impressive output.

Diffusion models are base of generative model. With its prominent contribution in Generative AI, text-to-image generation and LLMs, these models are trained to generate data close to data with which they are trained. The model is trained using 2 methods, Forward diffusion and Reverse diffusion.

These models work by destroying a data by continuously adding noise in Gaussian distribution to it. With each addition of noise, the data becomes more difficult to recognize until it becomes unrecognizable, that is a state of desired complex data distribution. This is called Forward diffusion or Diffusion process.

The reverse diffusion process involves recognizing the specific Gaussian noise patterns introduced at every step and training the model to denoise the data accordingly. The model uses its knowledge to predict the noise at each step and remove/denoise it. It may sound simple but involves complex reconstructions through Markov chain (A Markov chain is a stochastic model that uses mathematics to predict the probability of a sequence of events occurring based on the most recent event).

Once this model is trained, it is capable of generating data from pure noisy data. The result of training from forward and reverse diffusion results in a model capable of generating data from pure noise.

Diffusion model is highly used for generating outputs because of following reasons:

* High quality image generation with more realism and diversity.
* Can handle variety of input datatypes (text-to-image generation, layout-to-image generation)
* Can be trained with relatively less data.