**ADTA 5760: Midterm: PART III**

**Text-To-Images: Prompts & Responses: Analysis**

Professor: Thuan L Nguyen, PhD

Student Name: Azhan Saleem

# Question 1: What criterium/criteria can the student use to evaluate and critique the generative-AI performance of each LLM based on the paintings it has generated?

The student can use the following criteria to evaluate and critique the generative AI performance of each LLM:

**Realism**: Evaluate how realistic and natural looking the generated images are. Assess whether the AI successfully captures the intended realism, especially in the case of photorealistic prompts.

**Creativity:** Analyze the level of creativity exhibited by each LLM in interpreting and translating diverse prompts into visually interesting and novel images.

**Detail**: Examine the level of detail present in each generated image. Assess whether the AI models can produce intricate details, especially in complex scenes or close-ups.

**Consistency**: Evaluate the consistency of style across the three images generated by each LLM. Consistency is crucial for demonstrating the model's ability to maintain a particular visual theme or aesthetic.

# Question 2: Provide the reasons why the student has selected such property/properties as the criterium/criteria to evaluate and critique the generative-AI performance of each LLM?

**Realism**: Photorealism is often a key objective in generative AI applications, especially for prompts involving real-world subjects like faces. A good LLM should be capable of producing convincingly realistic images.

**Creativity**: Creativity is essential, especially for prompts that involve abstract or imaginative concepts like a mysterious bear surfing a cybernetic wave. A capable LLM should showcase its ability to think outside the box and generate visually interesting and innovative concepts.

**Detail**: The ability to capture intricate details is crucial for creating visually appealing and immersive images. A high level of detail indicates the model's capability to handle complex scenes and convey nuanced information.

**Consistency:** Consistency in style across different images demonstrates the robustness and reliability of the LLM. It shows that the model can maintain a coherent visual theme, which is essential for various applications, including artistic projects and branding.

# Prompt 1: Using the criterium/criteria, the student evaluates and critiques the generative-AI performance of each Text-To-Images LLM based on the images it has generated and write a report on the evaluation and critique.

## LLM: TTI\_LLM\_A

**Realism:**

TTI\_LLM\_A (img\_1\_A) generates an image that convincingly portrays a worried expression with high photorealism. The facial details, such as furrowed brows and nuanced expressions, contribute to the authenticity of the image.

**Creativity:**

The emphasis here is primarily on realism, and TTI\_LLM\_A adheres closely to this criterion. While the creative elements are restrained, the model effectively captures the essence of a worried expression without introducing unnecessary artistic liberties.

**Detail:**

The level of detail in the facial features is commendable. Fine lines, subtle expressions, and shading contribute to the overall detailed portrayal, enhancing the believability of the generated face.

**Consistency:**

TTI\_LLM\_A maintains consistency in style and detail throughout the image. The realistic approach is consistently applied to all facial features, creating a cohesive and authentic representation.

## LLM: TTI\_LLM\_B

**Realism:**

TTI\_LLM\_B (img\_1\_B) generates an image that captures a worried expression with a commendable degree of realism. The facial details are well-rendered, contributing to the overall authenticity of the image.

**Creativity:**

While realism is prioritized, TTI\_LLM\_B introduces subtle creative elements to the worried expression. The model manages to infuse a nuanced interpretation without compromising the primary objective of photorealism.

**Detail:**

TTI\_LLM\_B excels in detailing, portraying fine lines, facial contours, and subtle expressions effectively. The image reflects a careful consideration of intricate facial features, enhancing the overall quality of the generated face.

**Consistency:**

Consistency is maintained in terms of style and detail. TTI\_LLM\_B successfully blends realism with a touch of creativity, ensuring a coherent and consistent portrayal of the worried expression.

## LLM: TTI\_LLM\_C

**Realism:**

TTI\_LLM\_C produces an image (img\_1\_C) that excels in capturing a worried expression with an impressive level of realism. Facial details, expressions, and shading contribute to a highly convincing portrayal.

**Creativity:**

While prioritizing realism, TTI\_LLM\_C introduces subtle creative nuances in the worried expression. The model strikes a balance between authenticity and a nuanced creative interpretation, adding a layer of depth to the generated image.

**Detail:**

The level of detail is noteworthy, showcasing intricate facial features and expressions. TTI\_LLM\_C successfully conveys fine lines, textures, and shading, contributing to the overall quality and believability of the image.

**Consistency:**

Consistency is maintained across the image, reflecting a cohesive and realistic style. TTI\_LLM\_C ensures a harmonious blend of detail and creativity, resulting in a consistent portrayal of the worried expression.

# Prompt 2: Using the criterium/criteria, the student evaluates and critiques the generative-AI performance of each Text-To-Images LLM based on the images it has generated and write a report on the evaluation and critique.

## LLM: TTI\_LLM\_A

**Creativity and Concept:**

TTI\_LLM\_A (img\_2\_A) generates an illustration that successfully embodies the concept of a mysterious bear surfing a cybernetic wave. The bear's posture and the cybernetic wave are well-conceived, reflecting a creative interpretation of the prompt.

**Detail and Clarity:**

The level of detail in TTI\_LLM\_A's image is noteworthy. The intricate elements, such as the bear's fur, the cybernetic wave's components, and the overall composition, contribute to a visually engaging and clear illustration.

**Consistency in Style:**

TTI\_LLM\_A maintains consistency in style throughout the image, ensuring that the bear and the cybernetic wave blend seamlessly. The coherent style enhances the overall aesthetic appeal and makes the illustration visually appealing.

## LLM: TTI\_LLM\_B

**Creativity and Concept:**

TTI\_LLM\_B (img\_2\_B) generates an illustration that partially captures the concept of a mysterious bear surfing a cybernetic wave. While the bear is present, the cybernetic wave is less defined, impacting the overall realization of the prompt.

**Detail and Clarity:**

The level of detail in TTI\_LLM\_B's image is moderate. While the bear's features are discernible, the lack of clarity in defining the cybernetic wave affects the overall visual impact of the illustration.

**Consistency in Style:**

TTI\_LLM\_B maintains a consistent style in rendering the bear, but the less defined cybernetic wave introduces a slight inconsistency in the overall composition. This impacts the cohesiveness of the image.

## LLM: TTI\_LLM\_C

**Creativity and Concept:**

TTI\_LLM\_C (img\_2\_C) generates an illustration that deviates significantly from the prompt. The image does not effectively convey the concept of a mysterious bear surfing a cybernetic wave, introducing elements that diverge from the given criteria.

**Detail and Clarity:**

The level of detail in TTI\_LLM\_C's image is moderate. While the bear is discernible, the inclusion of unrelated elements diminishes the clarity of the overall composition.

**Consistency in Style:**

TTI\_LLM\_C introduces inconsistencies in style by incorporating elements that do not align with the prompt. The lack of cohesion in the representation impacts the overall visual appeal and cohesiveness of the illustration.

# Prompt 3: Using the criterium/criteria, the student evaluates and critiques the generative-AI performance of each Text-To-Images LLM based on the images it has generated and write a report on the evaluation and critique.

## LLM: TTI\_LLM\_A

**Creativity and Concept:**

TTI\_LLM\_A (img\_3\_A) generates a highly detailed photography close-up of a mysterious computer expert working on a laptop, with an FBI agent awaiting capture in the background. The model effectively captures the concept, creating a convincing portrayal of a scene from a sci-fi setting.

**Detail and Clarity**:

The level of detail in TTI\_LLM\_A's image is commendable. The intricacies of the computer expert's work on the laptop and the FBI agent in the background are well-realized, contributing to the overall clarity of the scene.

**Consistency in Style:**

TTI\_LLM\_A maintains consistency in style throughout the image, ensuring that both the computer expert and the FBI agent are seamlessly integrated into the scene. The coherent style enhances the overall photorealistic and intricate quality of the composition.

## LLM: TTI\_LLM\_B

**Creativity and Concept:**

TTI\_LLM\_B (img\_3\_B) generates an image that partially captures the concept of a highly detailed photography close-up of a mysterious computer expert. While the computer expert is depicted, the intricate details and the presence of an FBI agent are less pronounced.

**Detail and Clarity:**

The level of detail in TTI\_LLM\_B's image is moderate. The depiction of the computer expert is discernible, but the intricacies of the scene and the FBI agent lack the same clarity found in TTI\_LLM\_A's image.

**Consistency in Style**:

TTI\_LLM\_B introduces a slight inconsistency in style, particularly in integrating the computer expert and the FBI agent into the scene. The image lacks the same cohesiveness found in TTI\_LLM\_A's composition.

## LLM: TTI\_LLM\_C

**Creativity and Concept:**

TTI\_LLM\_C (img\_3\_C) generates an image that deviates significantly from the prompt. The scene depicted does not effectively convey the concept of a highly detailed photography close-up of a mysterious computer expert with an FBI agent awaiting capture.

**Detail and Clarity**:

The level of detail in TTI\_LLM\_C's image is limited. The representation of the computer expert lacks intricacy, and the presence of an FBI agent is unclear. The overall composition lacks the depth and clarity found in TTI\_LLM\_A's image.

**Consistency in Style:**

TTI\_LLM\_C introduces inconsistencies in style by deviating from the specified prompt. The lack of cohesion in the representation impacts the overall visual appeal and cohesiveness of the image.

# Prompt 4: Using the criterium/criteria, the student evaluates and critiques the generative-AI performance of each Text-To-Images LLM based on the images it has generated and write a report on the evaluation and critique.

## LLM: TTI\_LLM\_A

Creativity and Concept:

TTI\_LLM\_A (img\_4\_A) generates an image that partially captures the concept of a futuristic city with a neon sign saying, "EMERGE by Decrypt." However, the visibility and clarity of the text are limited, and the overall execution lacks a strong connection to the specified prompt.

Detail and Clarity:

The level of detail in TTI\_LLM\_A's image is moderate. While the futuristic cityscape is discernible, the neon sign's text, especially "Decrypt," is not clearly visible. The lack of clarity affects the overall impact of the composition.

Consistency in Style:

TTI\_LLM\_A introduces slight inconsistencies in style, particularly in rendering the neon sign. The lack of emphasis on the text affects the cohesiveness of the image.

## LLM: TTI\_LLM\_B

Creativity and Concept:

TTI\_LLM\_B (img\_4\_B) generates an image that partially captures the concept of a futuristic city with a neon sign saying, "EMERGE by Decrypt." The visibility of the text is improved compared to TTI\_LLM\_A, but the overall execution still lacks a strong connection to the specified prompt.

Detail and Clarity:

The level of detail in TTI\_LLM\_B's image is moderate. The futuristic cityscape is somewhat clear, and the neon sign's text, including "Decrypt," is visible but appears slightly blurry. The clarity of the text affects the overall quality of the composition.

Consistency in Style:

TTI\_LLM\_B introduces slight inconsistencies in style, particularly in rendering the neon sign. The blurriness of the text affects the cohesiveness of the image.

## LLM: TTI\_LLM\_C

**Creativity and Concept:**

TTI\_LLM\_C (img\_4\_C) generates an image that effectively captures the concept of a futuristic city with a neon sign saying, "EMERGE by Decrypt." The visibility and clarity of the text are notable, aligning well with the specified prompt.

**Detail and Clarity:**

The level of detail in TTI\_LLM\_C's image is commendable. The futuristic cityscape is clear and well-defined, and the neon sign's text, including "Decrypt," is visibly sharp. Clarity enhances the overall quality of the composition.

**Consistency in Style:**

TTI\_LLM\_C maintains consistency in style throughout the image, ensuring that both the cityscape and the neon sign are seamlessly integrated. The coherent style enhances the overall impact and cohesiveness of the composition.

# Conclusions: What conclusion can be made after analyzing the generative AI performance of each Text-To-Images LLM based on the images it has generated for each prompt mentioned above?

## LLM: TTI\_LLM\_A

TTI\_LLM\_A consistently demonstrates a decent but not outstanding performance in generating images based on the given prompts. While it generally captures the essence of the prompts, there are noticeable shortcomings in terms of detail, clarity, and adherence to specific instructions. TTI\_LLM\_A is a viable option for basic image generation but lacks the sophistication seen in superior models.

## LLM: TTI\_LLM\_B

TTI\_LLM\_B presents a mixed performance, showcasing improvement over TTI\_LLM\_A in some aspects. However, it struggles with maintaining a high level of detail, creativity, and precision in responding to the prompts. While it surpasses TTI\_LLM\_A in certain scenarios, it falls short of achieving a consistently high standard in image generation.

## LLM: TTI\_LLM\_C

TTI\_LLM\_C emerges as the most successful model in generating images that closely align with the provided prompts. It consistently exhibits a high level of creativity, detail, and adherence to specific instructions. TTI\_LLM\_C showcases a superior ability to understand nuanced prompts and translate them into visually compelling and accurate images. It outperforms both TTI\_LLM\_A and TTI\_LLM\_B across various scenarios.