Directorate of Online and Distance Education (DDOE)

SEMINAR-1

code : 23OMC107

Faculty Assigned: Dr. Umesh Tiwari
Student Name: Deepankar Sharma
Registration Number: 233512013
Course: MCA

Google Gemini

A Multimodal Revolution in AI

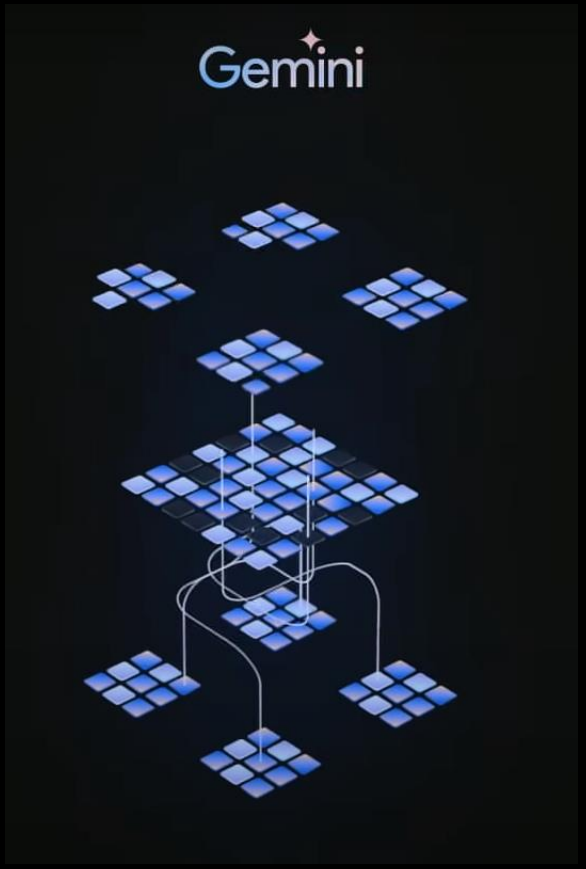
Introduction

- What is Gemini?

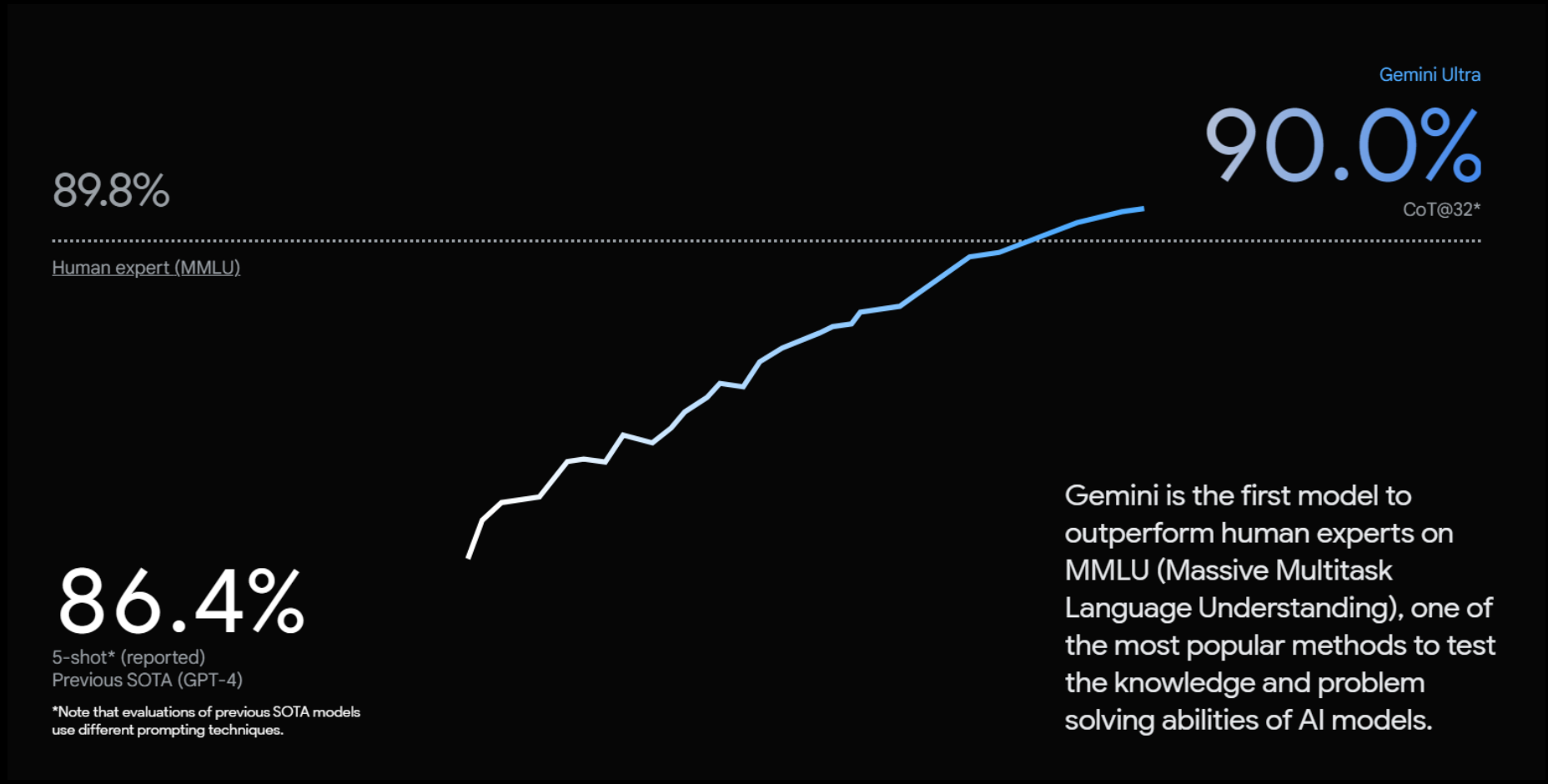
- Gemini is a family of multimodal large language models developed by Google DeepMind

- Gemini can understand and generate natural language across different modalities, such as text, speech, and vision

- Gemini is the successor to LaMDA and PaLM 2 (SOTA)



What’s special about Gemini?



Gemini Ultra, Pro, and Nano

- Three innovative models of Gemini, each tailored for different needs and scenarios

- Gemini Ultra: The ultimate solution for complex and demanding tasks, with the highest performance and capacity

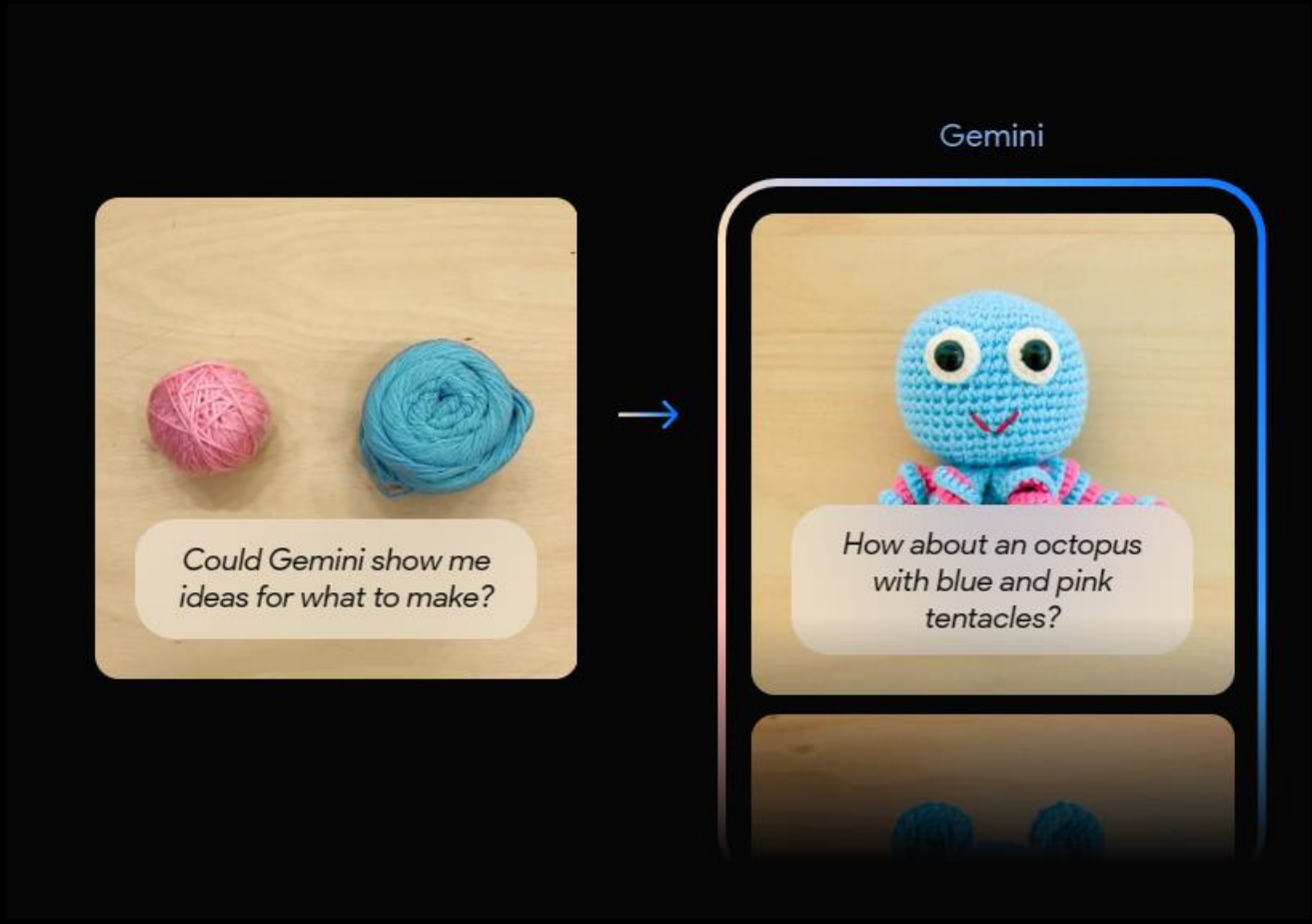
- Gemini Pro: The versatile solution for diverse and scalable tasks, with a balanced trade-off between power and efficiency

- Gemini Nano: The efficient solution for simple and on-device tasks, with a minimal footprint and low energy consumption

Multimodality At Core

- Gemini can understand, operate across, and combine various information formats

- This enables tasks like reasoning visually across languages



Beyond Text Generation

- Translation between languages.

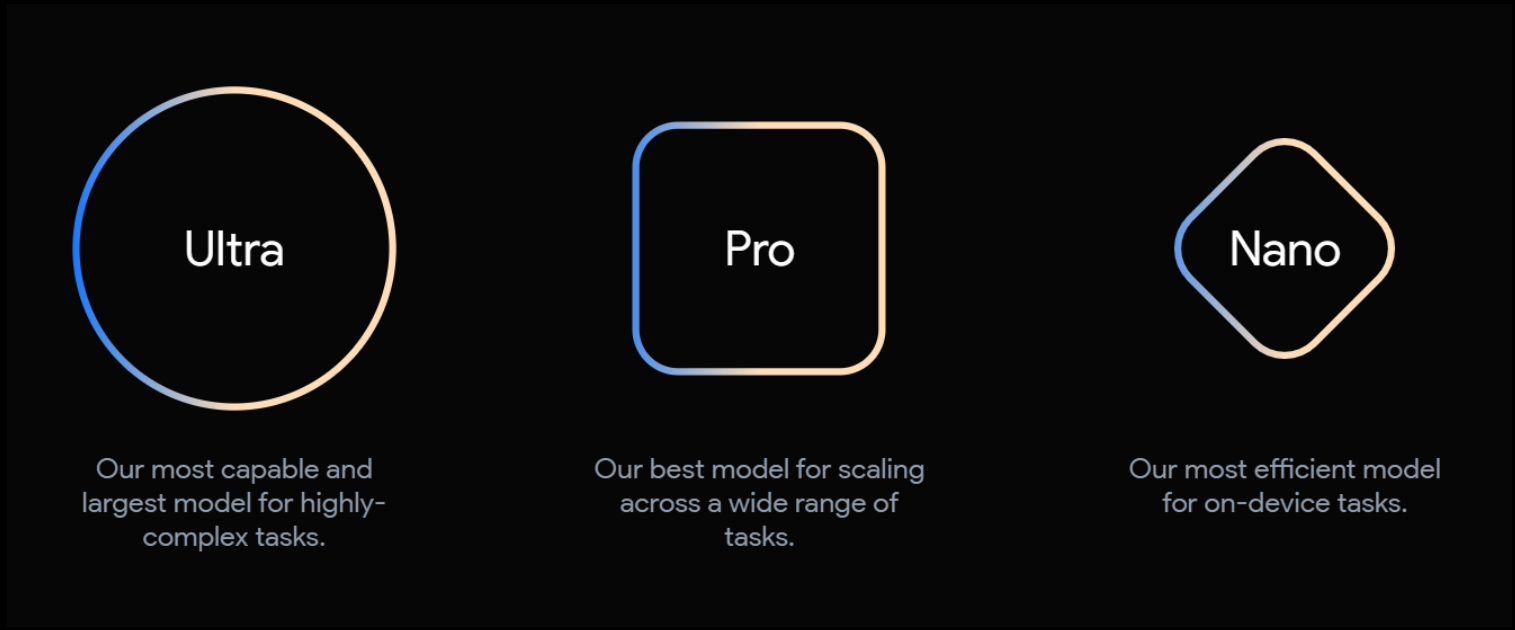
- Generating code based on natural language descriptions.

- Creating captions for images and videos.

- Summarizing video content.

How about a pig with blue ears?

Gemini: flexibility



- Supports various devices and platforms, from data centers to mobile phones

- Optimizes performance and resource utilization, adapting to different hardware capabilities and constraints

Responsible AI

- Google develops AI responsibly, with Gemini's comprehensive safety evaluations

- Gemini checks for bias and toxicity in AI models and outputs

- Google researches potential risk areas and applies rigorous testing techniques

Collaboration & Openness

- How Google leverages the collective wisdom of internal and external experts to develop Gemini

- How Google works with partners to ensure the robustness and fairness of Gemini models

The Road Ahead

- Google's plans for Gemini:

- Enhance capabilities to handle more complex tasks, such as natural language understanding, computer vision, and multimodal reasoning.

- Optimize efficiency and reduce resource requirements, such as memory, computation, and energy consumption.

- Address ethical considerations and ensure responsible use, such as privacy, fairness, accountability, and transparency.

Conclusion:

A New Era of AI

- Gemini: a landmark achievement in AI

- Understand the world through multimodal perception

- Interact with the world through natural language and actions

- Learn from the world through self-supervised and reinforcement learning

Thank You !