

# NOMAD Notes

Authors Mathematics and Computing  
Indian Institute of Technology

April 8, 2025

## 1 Introduction

### What is nomad?

NoMaD is a robot navigation system that can:

- Explore unknown places on its own (goal-agnostic behavior).
- Go to a specific place or object when given a goal image (goal-directed behavior).

The key idea: Instead of having two different systems for these tasks (like other methods do), NoMaD does both using one unified model.

### Why is this important?

In the real world:

- A robot might need to search for a place (like exploring a building).
  - Later, it may need to go back to that place (like retrieving an object from a room).
  - Existing systems often split these into two steps — exploration and navigation — with different models or planning stages.
- NoMaD combines both into one model, making the robot:
- Smarter.
  - Faster.
  - Less prone to mistakes (like collisions).

### How does it work?

Input: An image of the environment, and optionally a goal image (like “go here”).

Model: Uses a Transformer (like those in NLP/AI) to understand the visual world, and a Diffusion Model to plan actions.

Goal Masking: If a goal image is given, it uses that to guide movement. If not, it goes into exploration mode.

This ”masking” lets the model ignore the goal if it doesn’t have one, which allows exploration to happen naturally.

### **Why Diffusion Models?**

Diffusion models are great at generating complex, multi-modal outputs. That means:

The robot doesn't just plan a single path.

It can consider many possibilities, which helps it handle real-world messiness (like obstacles or uncertainty).