

# Immunova 2.0 - Functional Overview (Input/Output Perspective)

This document provides a detailed and accessible explanation of what each part of the Immunova 2.0 system does, based on how data flows into and out of it. This is intended for all team members, including those without deep technical or machine learning backgrounds.

---

## TIL CLASSIFICATION

-----

### Goal:

Classify tumor-infiltrating lymphocytes (TILs) in image patches from pathology slides.

### Inputs:

- Image patches from datasets/Non-TIL/ and datasets/TIL/
- Patch size typically 100-224px

### Outputs:

- Class: None, CD8+, CD4+, or Treg
- Optionally: Attention/Grad-CAM image in explainability/

### Core Files:

- til\_multiclass\_model.py: Basic ResNet classifier
- til\_graph\_model.py: Graph-based patch context classifier
- til\_cell\_segmented\_model.py: Combines segmentation and patch

---

## TREATMENT RESPONSE PREDICTION

-----

### Goal:

Predict whether a patient will respond to immunotherapy using both pathology images and gene expression.

Inputs:

- Image patch (from datasets/multimodal\_images/)
- Gene expression vector (from datasets/multimodal\_csv/gene\_expression.csv)
- Optional: cancer type label (domain ID)

Outputs:

- Binary classification: Response vs. No Response
- Log files, attention weights

Core Files:

- fusion\_model.py: Concatenation fusion model
- fusion\_attention\_model.py: Learns which modality to trust more
- domain\_adapted\_fusion\_model.py: Learns domain-specific embeddings
- immune\_signatures.py: Extracts only immune-escape-related genes

---

## SURVIVAL ANALYSIS

-----

Goal:

Estimate patient survival time and risk using molecular and TIL data.

Inputs:

- Clinical data (csv): survival time, event (death), optional biomarkers
- Time-series TIL data (csv): CD8+, CD4+, Treg levels over time

Outputs:

- Survival curves (Kaplan-Meier)
- Predicted risks for tumor progression, toxicity, and other outcomes

Core Files:

- survival\_transformer.py: Basic time-aware transformer

- km\_validation.py: Plots survival curve
- til\_survival\_multitask.py: Predicts multiple cause-specific survival risks

---

## DRUG OPTIMIZATION (REINFORCEMENT LEARNING)

-----

### Goal:

Recommend the best sequence of treatments for a patient, learning from historical success data.

### Inputs:

- Patient state (gene features, immune type, history)
- Available treatment actions (e.g., chemo, ICI, radiation)

### Outputs:

- Predicted optimal sequence of treatments
- Logs, reward plots

### Core Files:

- gnn\_model.py: Graph of treatment relationships
- rl\_optimizer.py: DQN-based agent
- ppo\_agent.py: Modern PPO agent with value-based learning

---

## SUPPORTING STRUCTURE

-----

- |                      |   |
|----------------------|---|
| datasets/            | -> All images and gene data             |
| explainability/      | -> Visual heatmaps, CAM overlays        |
| clinical_simulation/ | -> RL policy outcomes for patients      |
| checkpoints/         | -> Saved model weights (.pt files)      |
| docs/                | -> This document and architecture notes |

This structure allows the system to work end-to-end, from image to treatment simulation.