Toxicity prediction Master Thesis Proposal

This master's project proposes a progressive framework for toxicity prediction and functional subunit mapping in radiotherapy using machine learning (ML) and deep learning (DL). Radiotherapy efficacy is often limited by normal tissue toxicity, and current predictive models relying on dose–volume histograms fail to capture the spatial and biological complexity of tissue response. By integrating multimodal data, including clinical, baseline, and voxel-level dosimetric features, this work aims to develop interpretable, patient-specific models capable of identifying dose-sensitive subregions within organs at risk.

The database contains over 1300 patient data, already preprocessed and partially analyzed by two M2 students (O Ozer and M Harris)

The project will proceed in three phases. Phase 1A will focus on tabular data, starting with clinical variables alone, then incorporating baseline patient data and finally dosimetric parameters. LightGBM will serve as the primary model, and its performance will be evaluated using metrics such as AUC and calibration. Phase 1B will address DICOM-based modeling using voxel-level dose distributions. Multiple Instance Learning with Attention (MIL-Att) and Convolutional Neural Networks (CNNs) will be trained to predict toxicity patterns and to localize functional subunits associated with damage.

Phase 2 will integrate both modalities into a hybrid model combining MIL-Att and CNN architectures. The study will compare early versus late data fusion and test oversampling strategies to mitigate class imbalance. Model interpretability will be pursued through SHAP analyses for tabular data and Grad-CAM or attention maps for imaging, providing spatial insight into toxicity risk.

Expected outcomes include a benchmark of ML and DL models across data modalities, quantitative evaluation of the added predictive value of imaging, and spatial identification of radiosensitive regions. Clinically, this project aims to bridge voxel-level dose modeling and personalized risk prediction, supporting precision radiotherapy strategies that minimize toxicity while preserving tumor control.

Contact: Prof. Juliette Thariat LPC Caen and CFB Caen, thariat@lpccaen.in2p3.fr