# Generative flow models for modeling uncertainty in machine learning tasks

The modern machine learning models achieve incredible results in various tasks, including classification, regression, object detection, semantic segmentation, clustering, learning representation, generating multiple types of data, and many others. For the major of the mentioned applications, the goal is rather to provide the deterministic prediction, not focusing much on modeling the uncertainty of the predictions.

One of the most common applications of uncertainty modeling are areas of high-risk where the whole information about the predictive distribution is required, e.g., we would prefer to know that a patient have 50% chances to live next 12 months and 20% chances to live next 24 months, etc., rather than knowing only that mean time of survival for this patient is equal to 13 months. Another area of application is the object future state prediction, which is presented in Figure 1. As we can observe, such a problem requires a flexible approach enabling to model multiple modes of distribution.

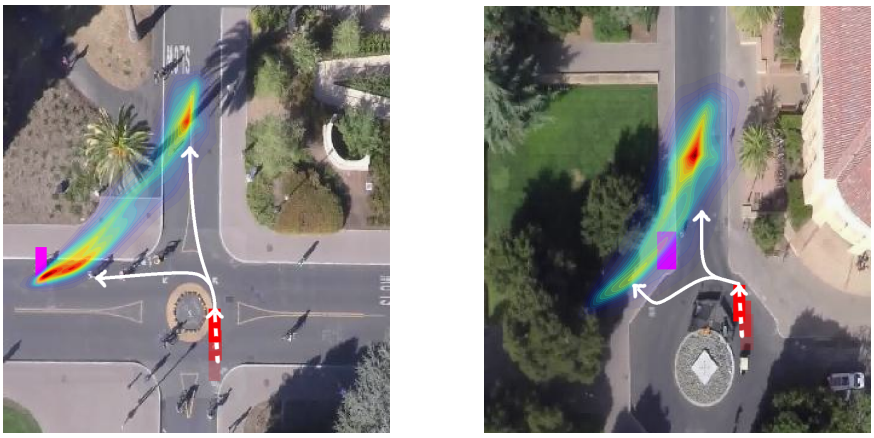

Figure 1: Application of one of the deep probabilistic regression models - RegFlow which utilizes flows. Given a past position of an object in image (in red) and training data, the goal of future prediction was to model a probability distribution of object future states (visualized with a heatmap) and the most probable directions (represented with arrows). Ground truth future states are colored violet.[1]

One of the main problems of the available probabilistic regression models is the assumption of the final probability distribution, which usually is a Gaussian distribution, due to the fact that it cannot model in a flexible way multimodal distributions. This limitation could be overcome by the flow generative models which are capable to model any probability distribution. Due to such a flexibility these models could be utilized not only in standard regression problems but also in more challenging setups.

The goal of this project is the construction of the deep learning flow-based models for univariate and multivariate probabilistic regression. Additionally, the flow models will be combined with tree-based models and extended to other machine learning tasks such as clustering or outlier detection, other types of data such as sequential data, and other areas of application. Moreover, a set of recommendations for flow model prior selection will be created and finally, the new tools for flow-based uncertainty analysis will be developed. Constructed methods will be analyzed qualitatively and quantitatively using the benchmark datasets publicly available on the Internet.

[1] *Source*: Maciej Zięba et al. *RegFlow: Probabilistic Flow-based Regression for Future Prediction.* (2020).