Intermediate Machine Learning: 1st lesson – Intermediate Machine Learning Basics

Welcome to Kaggle's Intermediate Machine Learning course!

If you have some background in machine learning and you'd like to learn how to quickly improve the quality of your models, you're in the right place! In this course, you will accelerate your machine learning expertise by learning how to:

* Tackle data types often found in real-world datasets (missing values, categorical variables)
* Design pipelines to improve the quality of your machine learning code
* Use advanced techniques for model validation (cross-validation)
* Build state-of-the-art models that are widely used to win Kaggle competitions (XGBoost)
* Avoid common and important data science mistakes (leakage)

Along the way, you'll apply your knowledge by completing a hands-on exercise with real-world data for each new topic. The hands-on exercises use data from the Housing Prices Competition for Kaggle Learn Users, where you'll use 79 different explanatory variables (such as the type of roof, number of bedrooms, and number of bathrooms) to predict home prices. You'll measure your progress by submitting predictions to this competition and watching your position rise on the leaderboard!

Prerequisites:

You're ready for this course if you've built a machine learning model before, and you're familiar with topics such as:

1} Model validation

2} Underfitting and overfitting

3} Random forests

If you're completely new to machine learning, please check out our Intro to Machine Learning course, which covers everything you need to prepare for this course.