The idea is that you use a linear model (that is defined by the features) and you subtract that from your predictions, therefore removing that signal from your predictions or in other words: neutralizing them against the features.

On youtube there is excellent material from arbitrage and also in the forum there is an in depth post on this.