

I tried this several years ago and found that the omission of certain data made my models worse. That said, I used completely different training data than what we are using now. I would be careful because I recall that my scores using this method were more volatile, not less. YMMV.