

## Abstrakt

Cílem této práce je za prvé poskytnout přesnější předpověď růstu HDP České republiky než oficiální odhad Českého statistického úřadu a České národní banky. Zadruhé rozšířit literaturu, zkoumající predikce časových řad s použitím strojového učení využívající data s různými frekvencemi. Přestože použité modely (Ridge model a Random Forest) nedokázaly překonat odhady oficiálních institucí, tato práce přispěla svými výsledky k rozšíření zatím málo prozkoumané oblasti zabývající se využitím strojového učení s daty o libovolných frekvencích. Vzhledem k tomu, že neexistuje model strojového učení, který by uměl pracovat s daty o různých frekvencích, tato práce ukazuje, jak proměnné transformovat do podoby vhodné pro jakýkoliv model. Dále je zkoumán efekt použití různých typu datasetů. Datasetsy se lišily v času předpovědi; konec současného čtvrtletí (nowcast) a 40 dní po referenčním čtvrtletí (backcast), typu transformace dataset; použití standardizovaných a nestandardizovaných dat a nakonec na nejlepším modelu (Ridge) je zkoumán vliv tzv. vysokofrekvenčních proměnných (na týdenní bázi). Zatímco u Random Forestu typ datasetu nehrál významnou roli, v případě Ridge modelu rozdílný dataset silně ovlivnil odhadované hodnoty. Hlavní rozdíl pak byl mezi netransformovaným a transformovaným datasetem, kdy při použití nestandardizovaného datasetu Ridge dosahoval celkově nejlepších výsledků, ale hlavně v post-Covidovém období byl velice volatilní. Z toho důvodu výběr konkrétního typu Ridge modelu bude záviset v jaké situaci bude model použit.

**Klasifikace JEL** C49, C32, C53, E66

**Klíčová slova** Predikce HDP, strojové učení

**Název práce** Predikce českého HDP pomocí strojového učení se smíšenou frekvencí