Tato práce se zabývá extrakcí informací z českých zpravodajských článků. Zaměřujeme se na čtyři úlohy: vydavatelský server, kategorie článku, textový gender autora a den vydání článk. Vzhledem k absenci vhodné datové sady pro tyto úlohy představujeme datovou sadu CZEch NEws Classification (CZE-NEC), jeden z největších českých klasifikačních datasetů, který je složen ze zpravodajských článků z různých zdrojů pokrývající období dvaceti let. Úlohy jsou řešeny pomocí Lineární regrese a předtrénovaných Transformerů. Důraz je kladen na metody dotrénování Transformerů, které jsou podrobně vyhodnoceny. Modely jsou porovnány s lidskými hodnotiteli, kteří zaostávají za modely na všech úlohách. Dále jsou modely porovnány s komerčním velkým jazykovým modelem GPT-3, který je překonán na polovině úloh, přestože je GPT-3 výrazně větší. Naše práce představuje silný startovní výsledek na sadě CZE-NEC, který umožňuje další výzkum v této oblasti.