V lingvistice se obvykle slova považují za složená z morfémů, což jsou dále nedělitelné jazykové jednotky nesoucí význam. Zadáním této práce je nalézt automatickou metodu dělení českých slov na morfémy, které by bylo možné přidat do DeriNetu, sítě derivačních vztahů mezi českými slovy.

Vytvořili jsme dvě různé takové metody. První nalézá hranice morfémů na základě hledání rozdílů mezi slovem a jeho derivačním předkem, a tranzitivně mezi všemi slovy v derivačním hnízdě. Tato metoda explicitně modeluje hláskové a morfologické alternace a nalézá nejvhodnější hranice morfémů pomocí metody maximální věrohodnosti. Ve srovnání s moderním systémem Morfessor FlatCat naše metoda přinejhorším mírně zaostává, ovšem v některých testech naopak dosahuje výsledků výrazně lepších.

Druhou metodou je neuronová síť pro současné předpovídání morfologické segmentace a derivačních předků, trénovaná na datech získaných první metodou a na derivačních vztazích ze sítě DeriNet. S naší hypotézou, že tento způsob trénování dvou úloh naráz pomůže k dosažení lepších výsledků oproti trénování samotné segmentace, jsou však ve shodě pouze některé provedené pokusy. Celkově dosahuje neuronová síť horších výsledků než první metoda, pravděpodobně kvůli trénování na datech obsahujících chyby, které se tím přidávají k chybám metody samotné.