

# SST annotálási útmutató

## Szentimentelemzés

Vadász Noémi és Ligeti-Nagy Noémi

2021. december 1.

### 1. Bevezetés

Az annotálási feladatok célja a Stanford Sentiment Treebank (SST) magyar nyelvűre történő átültetése. Az angol korpusz kb. 11000, az 1. példában láthatóhoz hasonló mondatot tartalmaz. Minden mondathoz tartozik egy szentimentcímke: pozitív vagy negatív.

- (1) With a cast that includes some of the top actors working in independent film, *Lovely & Amazing* involves us because it is so incisive, so bleakly amusing about how we go about our lives.

Az annotálás harmadik lépéseként a lefordított, szép magyar mondatokká alakított mondatokat kell azok érzelmi töltete alapján címkézni.

### 2. Az annotálás alapelvei

Az annotálás során egyszerre egy magyar mondatot látunk, amely eredeti angol formájában valamely filmről írt vélemény volt. A feladat annak megjelölése, hogy az adott mondat a filmről a) pozitív, b) negatív ítéletet fogalmaz meg, vagy c) semlegesnek tekinthető. A 2a példában egy semlegesnek tekinthető mondatot látunk, a 2b és 2c példák pedig egy pozitív, illetve negatív szentimentű mondatot tartalmaznak.

- (2) a. Bourne, Jason Bourne.  
b. Egy okos, édes és játékos romantikus vígjáték.  
c. Tényleg kell egy 77 perces film ahhoz, hogy megtudjuk, miért nem működik egy romantikus kapcsolat egy 15 éves fiú és egy 40 éves nő között?

Azt kell tehát eldönteni, hogy hogyan írható leglegjobban az adott értékelést író attitűdje, értékelése az értékelt film irányában:

- pozitív: a szöveg explicit vagy implicit jelét tartalmazza annak, hogy a mondat írója pozitívan értékelte a filmet: dicséri, ajánlja, elégedett volt vele
- negatív: a szöveg explicit vagy implicit jelét tartalmazza annak, hogy a mondat írója negatívan értékelte a filmet: mérges, csalódott, kritikus, szarkasztikusan nyilatkozik, gúnyolódik stb.
- semleges: a szöveg explicit vagy implicit jelét tartalmazza annak, hogy a mondat írója egyszerre pozitívan és negatívan is értékeli a filmet, vagy pedig a mondatban nincs semmilyen jele a mondat szerzője véleményének

### 3. A felület

Az annotálást az AT1 eszköz segítségével végezzük<sup>1</sup>. A nyitóoldalon bejelentkezés után az **sst\_sent** feladat kiválasztásával kezdheted a munkát.

Az annotálófelületen a felső sávban látható a bejelentkezett felhasználó neve, valamint a **Kijelentkezés** gomb. A felületen emellett látható az annotálandó mondat, valamint három gomb (**Negatív**, **Semleges** és **Pozitív**). A fenti elvek alapján a megfelelő gombra kattintással lehet címkézni a mondatot. Ekkor megjelenik egy visszajelző üzenet, ami a kiválasztott címkét mutatja. Ezen a ponton a mondatra vonatkozó döntés még bárhányszor megváltoztatható. A végleges döntés a **Következő mondat** gomb megnyomása után kerül rögzítésre.

A korábban mentésre került válaszok már nem módosíthatók, nincs mód a korábban annotált mondatokhoz történő visszatérésre, ezért kérek, minden esetben jól gondold meg a választ. A böngésző **Újratöltés**, **Vissza** ill. **Előre** gombjait kérlek, ne használd.

Az egyes mondatokhoz rendelt válasz tehát mindig a következő mondatra történő továbblépéskor rögzül, így a munkát bármikor félbe lehet hagyni. Ennek megfelelő módja a böngészőlap bezárása. Újbóli bejelentkezéskor az utoljára látott mondat jelenik meg, így ugyanonnan lehet folytatni a munkát.

Tájékoztató üzenet jelzi, ha elfogyott az annotálandó mondat (*Elfogyott a mondat!*).

Ha több különböző feladatban is részt veszel és át szeretnél térni másik feladatra, akkor se a böngésző **Vissza** gombjával próbálkozz, hanem lépj ki a **Kilépés** gombbal, majd jelentkezz be újra, hogy ki tudd választani a megfelelő feladatot a legördülő menüből.

---

<sup>1</sup><https://at1.nlp.nytud.hu/login>