

**Szegedi Tudományegyetem**  
**Informatikai Intézet**

**SZAKDOLGOZAT**

**Beszédes Bence**

**2025**

**Szegedi Tudományegyetem**  
**Informatikai Intézet**

**Teszteset generáló alkalmazás**

**Szakdolgozat**

**Szeged**

**2025**

<https://github.com/bempoj/testcase-generator>

## **Feladatkiírás**

A szakdolgozatom célja egy Python nyelven megvalósított alkalmazás létrehozása, amely képes egy projekt specifikációit feldolgozni, és egy előre meghatározott teszteset séma alapján automatikusan teszteseteket generálni.

Az alkalmazás biztosítja a projekt forráskódjából és kiadási jegyzeteiből származó információk integrálását, valamint az LLM-ek (nagy nyelvi modellek) támogatását a tesztesetek generálásában.

Az eredményként kapott adatok és tesztesetek legcélszerűbb tárolási módja a JSON formátum lenne, amik felhasználhatók, vagy beimportálhatók tesztelést rögzítő alkalmazásokba.

Az elkészült alkalmazás megkönnyíti és felgyorsítja a szoftverfejlesztéshez kapcsolódó tesztelési folyamatot.

## Motiváció

A szoftvertesztelői szakmában legtöbbször egy adott tesztelési ciklusban végrehajtott tesztesetek dokumentálása a legidőigényesebb feladat, ugyanis minden adatot (pl.: tesztlépéseket, elvárt eredményt) manuálisan kell leírnia a tesztelőnek.

Sokban megkönnyítené már akár a teszt tervezést is egy alkalmazás, amely egy projekt specifikációit, az adott funkció (feature) kódváltozásait és egy előre meghatározott teszteset sémát felhasználva a lehető legtöbb eshetőséget lefedő teszteseteket tudna generálni. A tesztelőnek csupán definiálnia kell egy általános struktúrát, ahogy szeretné, hogy a teszteset felépüljön, például az előfeltételekkel kezdve, a tesztlépéseket végrehajtva egy elvárt eredmény mellett, milyen jelenlegi működést tapasztal. Ezt a sémát és a feldolgozott specifikációkat felhasználva egy arra alkalmas nagy nyelvi modell (pl.: Claude AI) képes másodpercek alatt elég nagy lefedettséggel különféle teszteseteket generálni, megfelelő prompt megadásával.

## Nagy nyelvi modellek (Large Language Models)

A nagy nyelvi modellek olyan tanítható mesterséges intelligencia alapú modellek, amelyeket óriási mennyiségű adat feldolgozásával képeztek ki. A hatalmas adathalmaznak köszönhetően, látszatra képesek a természetes nyelv megértése, és összefüggő, értelmes mondatok generálására, illetve számos más tartalom előállítására (forráskód generálás, képgenerálás, hanganyag).

Miért fontosak az LLM-ek?

A mindennapi ember számára is elérhető LLM-ek váltak a generatív mesterséges intelligencia egyik legismertebb technológiájává, hiszen a széleskörű hozzáférés biztosította, hogy az AI bekerüljön a közéletbe. Környezetünk számos területén megtalálható manapság nagy nyelvi modellek által alkalmazva az AI. Annak ellenére, hogy az LLM-ek berobbanása egy hirtelen jelenségnek tűnhet, számos nagy vállalat éveket fordított ezen modellek implementálására és fejlesztésére.

A nagy nyelvi modellek kezdeti fázisukban úgynevezett alapmodellek (foundation models), amelyek képesek hatalmas adathalmazok betanulására mélytanulás és gépi tanulás révén.

Az utóbbi években a legnagyobb hangsúlyt a természetes nyelvfeldolgozás (NLP) és természetes nyelvi megértés (NLU) kapta. Emberek számára is a legmegfelelőbb és a legmegszokottabb módja az információátvitelnek a szöveges forma. Röviden fogalmazva, az LLM-ek úgy lettek kifejlesztve, hogy természetes nyelven megfogalmazott szöveget vagy akár más tartalmat generáljanak, a felhasználó által megadott adatok (input) alapján, amiből kikövetkeztetik, és szövegkörnyezetbe illesztik a választ (output). Képesek más nyelvre fordítani, összegezni szöveget és ezek mellett számítógépes kód generálásra. A hatalmas betanított adathalmaznak és a mérhetetlenül sok paraméternek köszönhetően különböző mintákat és kombinációkat rögzít, amiket felhasználva illeszti össze a generált eredményt.

Gépi tanulás – NLP, NLU

Claude, Anthropic API

Tesztelés fázisai – tesztesetek, tesztelési módszerek, eszközök

Python – Basic calculator app

JSON

Prompting

input, output