Kivonat

Az internetes böngészés rohamos terjedése és az egyre tágabb körű online ügyintézés ösztönöz, hogy felhasználói fiókokat hozzunk létre. Ezek által adataink biztonságosan eltárolódnak és több funkcióhoz is jutunk egy adott oldalon. Ahhoz, hogy ezekhez az adatokhoz más ne férhessen hozzá, különféle azonosítási módszereket dolgoztak ki a felhasználók védelme érdekében. Ennek ellenére, ha egy illetéktelen személy belépést nyer saját online fiókunkba, akkor nem csak felhasználói adatainkat, de akár anyagi javainkat is veszélyeztetheti.

Diplomamunkám célja egy olyan másodlagos hitelesítő rendszer létrehozása webes alkalmazások számára, amely a felhasználónév-jelszó párossal való beléptetés biztonságát növeli. Ezt úgy valósítsa meg, hogy összehasonlítsa a felhasználó által beírt karakterek leütési mintáját a már korábban bevitt mintákkal. Az alkalmazást egy PHP keretrendszerben, Laravelben készítettük el és az adatok tárolására a Firebase valós idejű adatbázist használtuk.

Az eredmény egy gyűjtő felület, ahol a kísérletben résztvevők néhány szöveg begépelésével egy billentyűzési adathalmazt szolgáltatnak. Ez által saját adathalmazra teszünk szert és kísérleteket hajthatunk rajtuk végre, amelyek hasznos eredményeket szolgáltathatnak a jövőbeli kutatásokra nézve. Emellett egy azonosító felületet is implementáltunk, ahol le tudjuk tesztelni éles helyzetekben is az alkalmazást.

***Kulcsszavak:*** billentyűzés alapú felismerés, másodlagos hitelesítés, Laravel, webes biztoság

Tartalomjegyzék

[1. Bevezető 2](#_Toc517887987)

[2. A projekt célja 3](#_Toc517887988)

[3. Bibliográfiai tanulmány 4](#_Toc517887989)

[4. Elméleti megalapozás 5](#_Toc517887990)

[4.1 A biometrián alapuló azonosítás áttekintése 5](#_Toc517887991)

[4.2 Billentyűzési biometria webes alkalmazásban 6](#_Toc517887992)

[4.2.1 Bemeneti mezők megválasztása 6](#_Toc517887993)

[4.2.2 Jellemzők kiemelése 7](#_Toc517887994)

[4.2.3 Az egyenletes hibaarány 7](#_Toc517887995)

[4.2.4 Felhasznált algoritmusok 8](#_Toc517887996)

[4.2.4.1 Manhattan scaled 8](#_Toc517887997)

[4.2.5 Kimenetek fúziója 9](#_Toc517887998)

[4.3 Felhasznált technológiák 9](#_Toc517887999)

[4.3.1 Adattárolás Firebaseben 9](#_Toc517888000)

[5. A rendszer specifikációi 10](#_Toc517888001)

[5.1 Követelmény specifikáció 10](#_Toc517888002)

[5.1.1 Kivonat 10](#_Toc517888003)

[5.1.2 Bevezető 10](#_Toc517888004)

[5.1.2.1 A dokumentum célja 10](#_Toc517888005)

[5.1.2.2 A dokumentum olvasói 11](#_Toc517888006)

[5.1.2.3 Rövid áttekintés 11](#_Toc517888007)

[5.1.3 Áttekintés 11](#_Toc517888008)

[5.1.3.1 Funkcionalitás 11](#_Toc517888009)

[5.1.3.2 Felhasználói osztályok és karakterisztikái 12](#_Toc517888010)

[5.1.3.3 Működési környezet 12](#_Toc517888011)

[5.1.3.4 Felhasználói segédletek 12](#_Toc517888012)

[5.1.3.5 Megszorítások és függőségek 12](#_Toc517888013)

[5.1.4 Rendszer követelmények 13](#_Toc517888014)

[5.1.5 Funkcionális követelmények 13](#_Toc517888015)

[5.1.6 Nem funkcionális követelmények 13](#_Toc517888016)

[5.1.6.1 Skálázhatóság és továbbfejleszthetőség 13](#_Toc517888017)

[5.1.6.2 Karbantarthatóság 13](#_Toc517888018)

[5.1.6.3 Nyelvi követelmények 13](#_Toc517888019)

[5.1.7 Interfész követelmények 13](#_Toc517888020)

[5.1.7.1 Felhasználói interfész 13](#_Toc517888021)

[5.1.7.2 Felhasználói interfész számítógépes eszközökhöz 13](#_Toc517888022)

[5.1.7.3 Felhasználói interfész mobil eszközökhöz 14](#_Toc517888023)

[6. Részletes tervezés 14](#_Toc517888024)

[6.1 Adatok tárolása Firebaseben 14](#_Toc517888025)

[6.2 Architektúra 14](#_Toc517888026)

[6.3 A rendszer tervezése és bemutatása 14](#_Toc517888027)

[6.3.1 A kontroller rész 14](#_Toc517888028)

[6.3.2 A modell rész 14](#_Toc517888029)

[6.3.3 A nézet (view) rész 14](#_Toc517888030)

[6.3.4 Segítő osztályok 14](#_Toc517888031)

[6.4 AJAX hívások bemutatása 14](#_Toc517888032)

[6.5 Adatok mozgása 14](#_Toc517888033)

[7. A rendszer felhasználása 14](#_Toc517888034)

[7.1 Adatgyűjtő felület 14](#_Toc517888035)

[7.2 Bejelentkezési felület 14](#_Toc517888036)

[8. Gyűjtő üzembe helyezése és kísérleti eredmények 14](#_Toc517888037)

[8.1 Gyűjtő üzembe helyezése 15](#_Toc517888038)

[8.2 Felmerült problémák és megoldásaik 15](#_Toc517888039)

[8.3 Kísérleti eredmények 15](#_Toc517888040)

[9. Következtetések 15](#_Toc517888041)

[9.1 Megvalósítások 15](#_Toc517888042)

[9.2 Összehasonlítás hasonló rendszerekkel 15](#_Toc517888043)

[9.3 További fejlesztési irányok 15](#_Toc517888044)

[10. Irodalomjegyzék 15](#_Toc517888045)

[11. Függelékek 15](#_Toc517888046)

# Bevezető

Az internet rohamos fejlődése által egyre több olyan tevékenységet vagyunk képesek lebonyolítani otthon, ami személyes jelenlétet igényelt volna évekkel ezelőtt. Manapság a legújabb okos telefonok megvásárlásától a banki átutalásokig bármit képesek vagyunk pár mozdulattal lebonyolítani és ezzel arányosan hatalmasra nőtt a felhasználói fiókjainknak az értéke is. Bárki, aki az adatainknak, vagy éppen azonosító tárgyainknak (pl. bankkártya, beléptető tóken) a tulajdonában van, képes arra, hogy hozzáférjen a személyes adatainkhoz, vagy akár az anyagi értékeinkhez is.

Tegyük fel, hogy nem biztonságosan kezeltük egy online banki alkalmazásnál használt fiókunkat és a belépési adatainkat megszerezte egy harmadik fél, akinek szándéka ezt a lehetőséget kihasználni. Ilyen esetben segít egy másodlagos hitelesítő rendszer, ami képes a detektálására az esetleges betolakodóknak. A rendszer észleli azt, hogy a belépett személynek egyes tulajdonságai nem egyeznek meg a regisztrált felhasználóéval, ezért jelzést küld és további megerősítést kér. Amennyiben a megerősítés nem megy végbe, a rendszer elkönyvelheti, hogy feltörés ment végbe és segítséget tud nyújtani a felhasználónak a további biztonsági lépések megtételéhez.

Egy ilyen rendszernél például ajánlatos viselkedési biometrián alapuló eszközökkel ellenőrizni, hogy valódi-e a felhasználó. Ezek az eszközök képesek arra, hogy a felhasználót a vele született, egyedi viselkedés mintái és mikro-mozzanatai alapján összehasonlítsák a korábbi belépéseihez és megállapítsák, hogy mekkora hasonlóság mérhető nála.

Általában egy ilyen rendszernél nem csak egyfajta viselkedést ellenőriznek, ugyanis a viselkedési biometrián alapuló hitelesítők nem képesek száz százalékos eredményt garantálni és megeshet, hogy a hamis felhasználót beengedik, vagy éppen a valós felhasználót kizárják. Ennek elkerülésére a rendszer több viselkedést is kell, hogy figyeljen és az egyes tulajdonságok (billentyűzés, egérmozgás, szóhasználat, stb.) mind visszaadnak egy szavazatot, amit összegezve megkapjuk, hogy mi a teendő az adott felhasználóval.

A fiókunk biztonsága érdekében tehát érdemes minél jobban megerősíteni a védelmet és erre egy módszer a billentyűzési ritmus figyelése. A felhasználónak tudnia nem kell róla, mert a háttérben működik. Emellett már egyre jobb és jobb eredményeket érnek el ilyen téren, tehát valószínű, hogy hamarosan egyre több oldalon lesz lehetőség ez a fajta másodlagos azonosítás is.

# A projekt célja

Az első fejezetben említett billentyűzés alapú azonosítási modul megvalósításában nyilvánul meg a célom. Ez magába foglalja egy gyűjtő rendszer létrehozását, mely segítségével a további kísérletekhez tudunk tanító és tesztelő adathalmazokat biztosítani, valamint egy azonosító rendszer létrehozását, ami segítségével a felhasználót egy második biztonsági réteggel is meg tudjuk védeni.

Jelenleg számos kísérlet folyik jobb és jobb eredményekért a billentyűzési ritmus alapú felismeréssel kapcsolatosan. Már a telegráf korában felfigyeltek az emberek a tényre, hogy be tudják egymást azonosítani a beütési ritmusuk alapján. Ez többnyire annak is köszönhet, hogy hogy az emberi agynak ugyanazok a neuro-fiziológiai részei kerülnek használatba gépeléskor, mint amik az aláírások esetében is. Bár egyre több kísérlet folyik ilyen irányban, teljes pontosságú megoldást még nem találtak és viszonylag kevés rendszernél alkalmazzák a módszert.

Az alkalmazás célközönsége az érzékeny adatokat kezelő alkalmazások köre, mivel ilyen esetekben kifejezetten fontos az, hogy az belépett felhasználó az legyen, aki a tulajdonosa a fióknak.

­Megvalósítás céljából a PHP egyik keretrendszerét, a Laravelt használtam. A választás azért esett a Laravel-re, mert számos ponton megkönnyíti a rendszer implementációját és amennyiben más PHP alapú rendszerbe szeretnénk a modult integrálni, kevés átalakítással képesek vagyunk rá. Az alkalmazás egyaránt képes gyűjteni adatokat a valós felhasználótól, valamint külön oldal van létre hozva arra is, hogy pár, a kísérletbe beavatott személy próbáljon belépni a tesztelt felhasználó adataival. Ezeket a belépési kísérleteket szintén elmentjük. A teljes adatbázis JSON formátumban van eltárolva a Firebase valós idejű adatbázisban, ezért a kísérletekhez, amiket Matlabban folytattunk, létrehoztunk egy átalakítót is, ami a JSON formátumban megadott adatainkat csv formátumba konvertálja.

Egy másik célja az államvizsga munkámnak, hogy minél hatékonyabb algoritmust, illetve jellemző kiemelő módszert találjak kísérletek által. Emellett az alkalmazást működővé tenni mobil eszközökön illetve táblagépeken is.

# Bibliográfiai tanulmány

# Elméleti megalapozás

## A biometrián alapuló azonosítás áttekintése

Az éppen bejelentkező felhasználóról eldönteni, hogy tényleg ő birtokolja-e az adott fiókot, több módon is lehetséges. Az azonosítási módszereket az alábbi három kategóriába soroljuk:

1. Tudás alapú azonosítás - Jelszó, minta illetve kérdés alapú ellenőrzések tartoznak ide. Legnagyobb előnyük hogy könnyű őket implementálni, míg hátrányuk, hogy el lehet őket felejteni, vagy éppen fel is törhetik pár algoritmus segítségével.
2. Token alapú azonosítás - Általában egy tárgy segítségével azonosítsák a felhasználót, ami csak az ő birtokában lehet. Ezt feltörni nehezebb és általában az elkészítése olcsó, de könnyű elveszíteni. Ide tartoznak a bankkártyák és a beléptető csipogók (például a Sapientia marosvásárhelyi székhelyének a bentlakásánál).
3. Biometrián alapuló azonosítás – A felhasználónak egy adott viselkedését vagy vele született jellemzőjét vizsgáljuk az azonosítás során. Két fajtája lehet:
   1. fiziológiai, amelyet nehéz hamisítani és magas pontossággal rendelkezik, mind például az ujjlenyomat vizsgálása vagy a retina ellenőrzése. Ezek a rendszerek viszont költségesek és mivel a felhasználó vele született fizikai tulajdonságát nézi, amennyiben egyszer sikerül megszereznie egy harmadik félnek a mintát, a felhasználó többé nem használhatja az adott módszert.
   2. viselkedési, melynél azt nézzük, hogy a felhasználó „hogyan” csinál valamit, nem pedig azt hogy mit csinál. Ilyen azonosításokat végezhetünk az egérmozgás, járás, hang, billentyű leütési ritmus és még sok egyéb egyedi mintát alkotó viselkedés elemzésével. Ezek a rendszerek általában nem biztosítanak teljes pontosságot, mivel megeshet, hogy kizárják a valódi felhasználót, vagy éppen beengedik a hamisat.

Bár a viselkedési biometrián alapuló rendszerek képesek arra, hogy egymagukban, viszonylag magas pontossággal azonosítsák a felhasználót, egy ilyen rendszer önmagában nem hatékony. Ezért is az alkalmazásnál ez a fajta hitelesítés nem veszi át a felhasználónév-jelszó páros szerepét, hanem az azonosított felhasználókat ellenőrzi le, hogy tényleg ők-e azok, akik eddig is folyamatosan bejelentkeztek.

Egy másik fontos része a billentyűzés alapú hitelesítésnek, mint bármely más viselkedésen alapuló biometriai hitelesítőnek, hogy egymagában mivel téves adatot ad, több ilyen ellenőrző modullal együtt képesek valós képet alkotni a felhasználó kilétéről. Például a billentyűzés mellett még lehet nézni az egér mozgását is. Ezért ajánlatos az alkalmazásban megírt azonosító modult egy szavazó komponensként alkalmazni az ellenőrzés során.

## Billentyűzési biometria webes alkalmazásban

Az dolgozatmunkám egy webes alkalmazásra épül, mivel sokkal könnyebben meg lehet valósítani az adatgyűjtést, ha egy szerveren rajta van az oldal és egyidejűleg korlátlan számú felhasználó írhatja be az adatokat. Ennek ellenére a felhasználónak lehetősége van arra, hogy ha nincs felügyelve, akkor egy rövid időre megszakítsa a gépelést és ez elrontsa az adott mintát. Emiatt párhuzamosan gyűjtöttem felügyelet nélküli beviteleket és általam felügyelteket is.

Az alkalmazás gyűjtésénél a JavaScriptben megírt JQuery könyvtár függvényeit használtam, melyek által le lehet kérni az adott billentyűnek a lenyomási és felengedési idejében kapott időbélyeget. Ezeket egy rejtett mezőben tárolva tovább lehet küldeni az űrlap beküldése során és a szerver feldolgozza, majd tárolja őket.

Egy másik előnye a webes alkalmazásoknak ilyen téren, hogy kevés módosítással képesek vagyunk arra, hogy a számítógép/laptop billentyűzetei mellett a rendszert működésre bírjuk okos telefonokon és táblagépeken is.

### Bemeneti mezők megválasztása

Tekintsünk most el pár pillanatra a programozás résztől és nézzünk rá a problémára úgy, hogy közben próbáljuk az emberi viselkedésnek a jellemzőit felhasználni.

Már számos adathalmazt hoztak létre, ezek közül én két ismertebbet próbáltam ötvözi, hogy a lehetséges jobb eredmény elérése mellett vizsgáljam meg azt is, mekkora különbség jelentkezik a felhasználó által megadott és a felhasználót a születésétől kísérő adatok beírásának hasznosságában billentyűzési ritmus alapú ellenőrzésnél. A két billentyűzés dinamikai adatbázis a következő:

* ATVS-Keystroke database [1] – 63 felhasználónak 12 valós és 12 imposztor általi belépése van tárolva. Minden belépés során a felhasználók a személyes adataikat adták meg, amik a vezetéknév, keresztnév, e-mail cím, nemzetiség és személyi igazolvány szám. Két szesszióban folyt le a valós felhasználótól az adatgyűjtés, szessziónként 6 belépéssel. Az adathalmaz előnye, hogy a felhasználó egy olyan szöveget kell, hogy beírjon, amit már jól ismer és ezért összefüggőbbek lesznek a mintái.
* CMU database [2] – 51 felhasználó 8 szesszióban adta le az adatait és két szesszió között legalább egy nap telt el. Ennél az adathalmaznál a felhasználók csak egy jelszavat gépeltek be, ami mindenkinek előre meghatározott volt és tartalmaz kis- és nagybetűt, speciális karaktert valamint számot. Az adathalmaz előnye, hogy nem volt szűkség külön imposztorokra, mert mindenki ugyanazt a jelszavat ütötte be.

Ismerve a fent említett két adathalmazt, olyan bemeneti mezőket választottam, amik ötvözik ezeknek a tulajdonságait és általuk olyan kísérleteket hajthatunk végre, amiket külön a két adathalmazon nem tehettünk volna. Ehhez az első mezőnek a *felhasználónevet*, a másodiknak az oldalon ügyködő személynek a *valódi nevét*, míg a harmadik mezőnek egy *training password* nevű adatot kellett tartalmaznia. Ez utóbbi kötelező módon kell tartalmazzon kis- és nagybetűt, számot és speciális karaktert.

A felmerült kérdések a következők voltak: Ha a legszemélyesebb a felhasználónév akkor az is lesz a leghatékonyabb azonosításnál? Amennyiben igen, a jelszó, amit nem látunk milyen mértékben lesz kevésbé hasznos az azonosításnál, mint a másik két mező. Ezeknek a megválaszolását a 8.3-as szekcióban lehet megtalálni a többi kísérleti eredmény mellett.

### Jellemzők kiemelése

Az adatok bevitele után következik a rendszer tanításának első része, ami a jellemzők kiemeléséből áll.

Amikor az adatok megérkeznek a szerverre, még nyers formában vannak, és úgy lehet őket kezelni, mint egy lista időbélyeg párosokból. Ezeket az adatokat nem lehetne felhasználni azonosításra, mert az időbélyegek folyamatosan változnak. Ezért a beérkező billentyűk leütési és felengedési időpontjait az őket követő billentyű leütési időiből kivonjuk. Ez által egy úgynevezett di-gráfot hozunk létre, mert kettőnként csoportosítva nyerjük ki az adatokat, amiket felhasználunk a minta elkészítésére, ami által később a felhasználót tudjuk azonosítani.

A tulajdonságok, amiket jelenleg kinyerünk a következők:

* Két egymás utáni leütés közti idő (DD = Down-Down)
* Két egymás utáni felengedés közti idő (UU = Up-Up)
* Az egyik billentyű leütése és a következő felengedése közti idő (DU = Down-Up)
* Az egyik billentyű felengedése és a következő leütése közti idő (UD = Up-Down)
* Az egyik billentyű lenyomása és felengedése között eltelt idő (H = Hold)

[rajz]

Ezeket a tulajdonságokat még kibővíthetjük, hogy hármanként is nézze az időzítéseket, viszont ez az információ ugyanaz, mintha összegeznénk két di-gráf elemet, ezért nem is emeltem ki több jellemzőt.

### Az egyenletes hibaarány

A fent említett finomított adatokat átadhatjuk egy algoritmusnak, ami a szívét fogja képezni a rendszerünknek. Egy jó algoritmus nagyságrendekkel növelheti az elért eredményt, ami alatt minél több elfogadott valódi felhasználót és elutasított hamisat értünk. Az algoritmusok jóságát fontos hogy tudjuk összemérni, amire az úgynevezett Egyenletes Hibaarányt (EER = Equal Error Rate) használjuk. [ref]

Minden kiértékelt bejelentkezési próbálkozás kap egy pontszámot, amiről a vágáspont alapján eldönti, hogy valós felhasználónak könyveli-e el az aktuálisat, vagy hamisnak. Az EER-t két tényező határozza meg: a hibásan elfogadott hamis felhasználók aránya és a hibásan elutasított valós felhasználóké. Ezt legegyszerűbben úgy tudjuk vizualizálni, ha az adott belépési kísérleteket ábrázoljuk az általuk szerzett pontszám függvényében. Ezt elvégezve a valós és hamis felhasználókra ugyanazon az ábrán két görbét kapunk, melyeknek a metszéspontja adja az Egyenletes Hibaarányt.

[ábra]

Az EER értéke minél kisebb, az algoritmus annál jobban működik. Napjainkban már vannak algoritmusok, amik közel 0% EER-t érnek el, viszont nagyon sok tanító mintát igényelnek és ezért valós rendszereknél nem működnének.

Mikor küszöbértéket, vagy más néven vágáspontot választunk, egy olyan függőleges egyenest helyezünk el a grafikonon, ami vagy az EER ponton helyezkedik el, vagy abban irányba mozog, amerre több hamis felhasználót ki tud zárni, ahogy az az ábrán (ábra) is látszik. Minden következő belépési kísérlet pontszáma akkor lesz elfogadva, ha meghaladja ezt az értéket. Ez lehet minden felhasználóra egy globális érték, vagy felhasználónkét váltakozó.

Valós rendszereknél nem áll rendelkezésre az imposztor adathalmaz, ezért a vágáspontot más módon kell, hogy beállítjuk [ref]. Erre az egyik megoldás az, hogy kizárjuk a felhasználó mintáinak leggyengébb 10 százalékát. Ezzel általában elég magasra tudjuk helyezni a vágáspontot ahhoz, hogy kizárjuk a lehetséges imposztorok nagy részét. (ábra)

### Felhasznált algoritmusok

A következőben tárgyalt algoritmusokat először Matlab alatt teszteltük le a CMU [2] adathalmazon. Tesztelés után integráltuk az algoritmusokat a webes alkalmazásunkba is.

#### Manhattan scaled

Az első algoritmus, amit kipróbáltunk, mivel egyszerűsége ellenére az egy nagyon jó eredményt ad az eddigi felmérések szerint [3][ref]. Az említett adathalmazon tesztelve 0.09% EER-t kaptunk, amire a vágáspont 0.047. Az algoritmus a tulajdonságok közti távolságon alapszik [4], és egy átlagolásból, valamint egy standard eltérés számolásból áll.[ábra]

Első lépésként nyers adatokból kinyert tulajdonságokat (DD, UU, DU, UD, H) sorra átlagoljuk a következő képlet szerint:

(1) Ahol az tulajdonság átlagát, a minták számát, és a az tulajdonság mintáját jelenti.

Eredményképpen egy vektort kapunk, ami sorra tartalmazza az összes leütött karakter tulajdonságainak az átlagát. Következőnek standard eltérést számolunk a következőképpen:

(2) Ahol az tulajdonság standard eltérését, a minták számát, és a az tulajdonság mintáját jelenti, valamint az tulajdonság átlagát.

Ez szintén egy vektort fog adni, ami a tulajdonságok eltérését mutatja az átlagtól. Ezt az értéket a pontosabb eredmény érdekében minden tulajdonságnál leskáláztuk –el. Az átlag és standard eltérés segítségével már ki tudjuk számolni egy megadott minta pontosságát, ami egy pontszám lesz. Hogy a pontszámot megkapjuk, a következő képletet alkalmaztuk:

(3) Ahol a végső pontszámot jelenti.

A fenti képlet alapján minél közelebb van egy minta, annál kisebb lesz a végső pontszám [ábra]. Hogy 0 és 1 értékek közé skálázzuk, és megfordítsuk a pontszám növekedését, hogy a közeli minták nagyobb pontokat adjanak, a (4) képletet használtuk. A felhasználónkénti vágáspont meghatározásáért kizártuk a legkisebb pontszámot elért 10%-ot.

### Kimenetek fúziója

## Felhasznált technológiák

### Adattárolás Firebaseben

Választásunk azért esett a Firebase felhő alapú, valós idejű adatbázisra, mert gyorsan le lehet kérni belőle az adatokat, könnyű szemmel tartani a változásokat, valamint a JSON típusú adatok tárolása helyzetünkben sokkal előnyösebb, mint egy relációs adatbázisban táblákat használni.

Direkt Laravel támogatást a Firebase nem nyújt, ezért egy külön modul segítségével oldottuk ezt meg, amit GitHub-on publikáltak. Miután hozzáadtuk a Laravel projektünkhöz az adott modult, egy Firebase által generált kulcs JSON file segítségével bárhol le tudtuk kérni az adatbázis referenciát. Miután a referenciát megkaptuk, a következő műveleteket lehet végre hajtani:

* getReference(útvonal)
  + Visszatérít az adott útvonalra egy hivatkozást.
* orderByChild(csomópont neve).
  + Az adott csomópont szerint fogja rendezni a lekért adatokat, amennyiben a csomópont létezik.
* equalTo(érték)
  + Amennyiben egy adott érték szerint már rendezve van az orderByChild tagfüggvény által a lekérendő érték, ez a függvény csak azokat a csomópontokat kéri le, amelyek egyenlők a megadott értékkel.
* getValue()
  + Visszatéríti a kért csomópontokat.
* set(érték)
  + Beállítja az adott referenciára a megadott értéket. Ha egy asszociatív tömböt adunk meg akkor az elemek egymás alatt, hierarchikusan fognak elhelyezkedni.

A következőkben néhány példát láthatunk az itt megemlített függvények működésére: [példák: kép–nodark + szöveg]

# A rendszer specifikációi

## Követelmény specifikáció

### Kivonat

A dokumentum tartalmazza azon követelményeknek leírását, melyek betartásával az alkalmazást sikeresen el lehet készíteni. A specifikáció tartalmazza az alkalmazásra vonatkozó követelményeket, mind a gyűjtő, mind a tesztelő részére vonatkozólag.

A dokumentum első szekciója a bevezető, amely általánosan bemutatja a dokumentum célját, valamint rövid áttekintést nyújt arról. Ezt követi az alkalmazás részletes áttekintése, ami során említést teszünk a funkcionalitásokról, felhasználói osztályokról, függőségekről és megszorításokról. Ezek után a követelmények lesznek tárgyalva, és minden funkcionalitás részletes leírása.

### Bevezető

#### A dokumentum célja

A dokumentum tartalmazza a billentyűzési adatokat gyűjtő és ezen adatok alapján azonosító alkalmazás specifikációját. A cél a követelmények részletet és egyértelmű megfogalmazása. Ezek a követelmények a projekt megtervezéséhez szükségesek, és érvényesülniük kell a végleges termékre.

#### A dokumentum olvasói

A dokumentum célközönsége magába foglalja:

* Az alkalmazás fejlesztőit
* Az alkalmazás továbbfejlesztőit
* Az alkalmazás tesztelőit

#### Rövid áttekintés

A projekt célja egy olyan webes alkalmazás létrehozása, ami képes egy felhasználót hitelesíteni a bejelentkezési mezőkbe írt karakterek leütési ritmusa alapján. A hitelesítés mellett egy olyan lehetőséggel is kell, hogy rendelkezzen az alkalmazás, mely biztosít egy könnyen kezelhető adatgyűjtő felületet. A weboldal külsejénél fontos az esztétikus megszerkesztés, valamint a dinamikus navigálási lehetőség.

A projekt két részre bontható: először el kell készíteni a bejelentkezési felületet, majd a JavaScript gyűjtőt, ami által kinyerjük a felhasználó billentyű leütési időbélyegjeit.

### Áttekintés

#### Funkcionalitás

A funkcionalitások a bejelentkezés illetve a tanítás köré csoportosulnak leginkább. Emellett fontos, hogy a felhasználó tudjon regisztrálni az oldalra, az elfelejtett jelszavát tudja újra beállítani e-mail cím által, valamint tudja megtekinteni tanítás esetében a saját folyamat állapotát.

A projektre vonatkozó funkcionalitások listája a következő:

* Regisztrálási funkció
  + A szükséges adatok megadásával felhasználói fiók létrehozása
* Bejelentkezési funkció
  + Valós felhasználónév-jelszó kombinációval belépés a tanító oldalra
  + Tesztelés esetében, ha a bejelentkező személy átmegy a másodlagos ellenőrzésen is, megjeleníteni, hogy sikeresen bejelentkezett
* Jelszó újraállítási funkció
  + A felhasználó az e-mail címét megadva egy link által újra tudja állítani a bejelentkező jelszavát
* Tanítási funkció
  + A felhasználó beírt adatainak és rögzített billentyűzési adatainak a lementése, amennyiben azok valósnak bizonyulnak
  + A ritmust lehetségesen elrontó karakterek letiltása vagy figyelmen kívül hagyása, helyzettől függően
  + Letiltott karakterek esetén hibaüzenet megjelenítése
* Folyamat megtekintési funkció
  + A felhasználó kell, hogy lássa, hány helyes mintát küldött be, mennyi kell egy szesszió befejezéséhez, és hány szessziót vitt végig
* Nyelvi funkció
  + Angol és magyar nyelvű felület
* Segítő funkció
  + A választott nyelven visszaküldött üzenet érvénytelen adatok esetén
  + Leírása a kísérletnek, a letiltott és ignorált karaktereknek, valamint az egyéb lehetséges félreértések tisztázása, mind például a tanító jelszó és a bejelentkezési jelszó közti különbség
  + Megjelenítése tanító módban a beírandó adatoknak

#### Felhasználói osztályok és karakterisztikái

Figyelembe véve, hogy a felhasználók lehetnek tesztelők, vagy a gyűjtésben résztvevők, akik a tanító halmazt bővítik, a következő két csoportba sorolhatjuk őket:

* Tesztelő felhasználók
  + Leginkább használt funkcionalitások, sorrendben:
    - Regisztrálási funkció
    - Bejelentkezési funkció
    - Segítő funkció
    - Jelszó újraállítási funkció
    - Nyelvi funkció
* Tanító felhasználók
  + Leginkább használt funkcionalitások, sorrendben:
    - Tanítási funkció
    - Bejelentkezési funkció
    - Regisztrálási funkció
    - Segítő funkció
    - Folyamat megtekintési funkció
    - Jelszó újraállítási funkció
    - Nyelvi funkció

#### Működési környezet

Az alkalmazás egy internetes böngésző segítségével érhető el a kliens számára.

#### Felhasználói segédletek

A segítő funkció által a felhasználó rálátást nyer arra, milyen lehetőségei és megkötései vannak az oldal használata során. A regisztrálás nézeten belül specifikálva van, hogy milyen karaktereket használhat a felhasználó egyes mezőknél, valamint tisztázva vannak a mezők jelentései.

A tanító oldalon minden bejelentkezés során megjelenik egy jelzés, ami figyelmezteti a felhasználót arra, milyen karaktereket üthet le, valamint milyen megszorításokhoz kell, hogy tartsa magát az adatbevitel alatt illetve milyen segítséget nyújt az oldal. Az említett oldalon szintén lehetősége van arra a felhasználónak, hogy megtekintse a beírandó adatokat.

#### Megszorítások és függőségek

Az alkalmazás használatához első sorban az internetes kapcsolat elengedhetetlen.

Ezen kívül ajánlott minél újabb böngészőt használni. A böngészőnél be kell, hogy legyen kapcsolva a JavaScript, különben sem a billentyű leütési időbélyegek lekérése és továbbküldése, sem a többi JavaScript alapú funkció nem lesz elérhető.

### Rendszer követelmények

Az alkalmazást egy szerverre kell feltölteni, hogy elérhető legyen az interneten. Szükséges egy web szerver, ajánlott az Apache. A Laravel 5.5 verziója igényeli a PHP 7.2 verzióját, valamint a Composert. Adatbázis konfigurációra nincs szűkség, mivel az összeköttetés a kulcs segítségével megoldódik.

### Funkcionális követelmények

5.1. Táblázat - Regisztrálási funkció

|  |  |
| --- | --- |
| **Regisztrálási Funkció** | |
| **Leírás** | A szükséges adatok megadásával felhasználói fiók létrehozása |
| **Prioritás** | Nagyon fontos |
| **Kiváltás** | 1. A felhasználó a bejelentkezési oldalon rákattint a *Fiók regisztrálása* linkre. |
| **Funkcionális követelmény** | A regisztrációs oldalra való navigálás után a felhasználó ki kell, hogy töltse a következő mezőket:   * Teljes név – szöveg, legalább 8 karakter * Életkor – szám 10 és 120 között * Felhasználónév – szöveg, legalább 8 karakter * E-mail cím – valós e-mail cím formátum * Tanítandó jelszó – ezt fogja a felhasználó a többször bevinni, legalább 8 karakteres szöveg, mely tartalmaz kis- és nagybetűt, számot és speciális karaktert * Nemzetiség – legördülő menü, lehet: Román, Magyar, Angol * Jelszó – a belépéshez van rá szűkség, rejtett szöveg, legalább 10 karakteres * Jelszó megerősítés – rejtett szöveg * Nem: legördülő menü, lehet: Férfi, Nő * Domináns kéz: legördülő menü, lehet: Bal, Jobb * Eszköz: legördülő menü, lehet: Laptop, PC, Táblagép, Okos telefon (az utolsó kettő le van tiltva, még nem működőképes) * Szesszió kód - ellenőrzött adatbevitel esetére, ha nem ellenőrzött maradhat üresen   Emellett fontos, hogy a felhasználó elfogadja a Felhasználói Feltételeket.  Amennyiben minden mező helyesen ki van töltve és a felhasználó elküldte az űrlapot, átirányítódik a bejelentkezés oldalra. Hiba esetén az alkalmazás jelez, hogy melyik mezőben, és milyen típusú hiba jelent meg. |
| **Megszorítás** | - |

5.2. Táblázat - Bejelentkezési funkció

|  |  |
| --- | --- |
| **Bejelentkezési Funkció** | |
| **Leírás** | Valós felhasználónév-jelszó kombinációval belépés a tanító oldalra.  Tesztelés esetében, ha a bejelentkező személy átmegy a másodlagos ellenőrzésen is, megjeleníteni, hogy sikeresen bejelentkezett. |
| **Prioritás** | Nagyon fontos |
| **Kiváltás** | 1. A felhasználó a főoldalra navigál. 2. A felhasználó nincs bejelentkezve, de olyan útvonalat próbál elérni, amit csak bejelentkezett felhasználók használhatnak. 3. A regisztrációs oldalon a felhasználó a *Vissza a bejelentkezéshez* linkre kattint. 4. A felhasználó kijelentkezik. |
| **Funkcionális követelmény** | A felhasználónak rendelkeznie kell egy valós fiókkal. Felhasználónév-jelszó párossal tud regisztrálni.  A felhasználó amennyiben adatokat akar bevinni az adathalmaz bővítéséhez, a billentyűzési stílusát ellenőrző algoritmus is érvénybe lép, és amennyiben a mintája nem elég pontos, vissza kerül a főoldalra hibával.  Sikeres bejelentkezés esetén:   * Tanító módban: A felhasználó a tanítást lebonyolító oldalra lesz átirányítva. * Tesztelő módban: A felhasználó vissza lesz küldve a bejelentkezési oldalra egy sikeres belépést közlő üzenettel. |
| **Megszorítás** | Amennyiben a felhasználó teszteli a rendszert:  A böngészőben be van kapcsolva a JavaScript. |

5.3. Táblázat - Tanítási funkció

|  |  |
| --- | --- |
| **Tanítási Funkció** | |
| **Leírás** | A felhasználó beírt adatainak és rögzített billentyűzési adatainak a lementése, amennyiben azok valósnak bizonyulnak.  A ritmust lehetségesen elrontó karakterek letiltása vagy figyelmen kívül hagyása, helyzettől függően.  Letiltott karakterek esetén hibaüzenet megjelenítése. |
| **Prioritás** | Nagyon fontos |
| **Kiváltás** | 1. A felhasználó sikeresen bejelentkezett |
| **Funkcionális követelmény** | A felhasználó három mezőbe kell, hogy tudjon írni adatokat:   1. Felhasználónév – szöveg, legalább 8 karakter 2. Teljes név – szöveg, legalább 8 karakter 3. Tanítandó jelszó - legalább 8 karakteres szöveg, mely tartalmaz kis- és nagybetűt, számot és speciális karaktert   Amennyiben a felhasználó a letiltott billentyűk közül valamelyiket lenyomja, a mező üresre állítódik, nullázódik a billentyűzési adat az adott mezőre, és megjelenik egy hibaüzenet 5 másodpercre, majd eltűnik.  A letiltott billentyűk a következőek: *Backspace, Delete, Shift/Alt* lenyomása majd felengedése, közben leütött karakter nélkül, illetve *CapsLock* lenyomása kétszer, anélkül hogy közben karakterek is leütődtek volna.  A *Tab, Enter, Ctrl* és navigáló billentyűk figyelmen kívül vannak hagyva billentyűleütési szempontból.  Ha a felhasználó három másodpercen keresztül nem ír egy adott mezőbe, akkor az ezt követő karakter beírás lenullázza a mezőt, így megoldva azt, hogy a felhasználónak, ha elterelődik a figyelme, ne romoljon az adat.  Sikeres adatbevitel esetén az oldalon megjelenítődik, hogy egyel több minta van leadva, a mezők újra állítódnak, és a felhasználó beviheti a következő adatot, |
| **Megszorítás** | 1. A böngészőben be van kapcsolva a JavaScript. |

5.4. Táblázat - Jelszó újraállítási funkció

|  |  |
| --- | --- |
| **Jelszó Újraállítási Funkció** | |
| **Leírás** | A felhasználó az e-mail címét megadva egy link által újra tudja állítani a bejelentkező jelszavát. |
| **Prioritás** | Közepesen fontos |
| **Kiváltás** | 1. A bejelentkezés oldalon az *elfelejtett jelszó* linkre kattintva. |
| **Funkcionális követelmény** | A felhasználónak kell rendelkeznie egy már meglévő fiókkal, és a fiókhoz tartozó e-mail címnek valódinak kell lennie. Ez után az adott e-mail címre elküldődik egy link, amit követve a felhasználó újra be tudja állítani a jelszavát. |
| **Megszorítás** | 1. A böngészőben be van kapcsolva a JavaScript. |

5.5. Táblázat - Folyamat megtekintési funkció

|  |  |
| --- | --- |
| **Folyamat Megtekintési Funkció** | |
| **Leírás** | A felhasználó kell, hogy lássa, hány helyes mintát küldött be, mennyi kell egy szesszió befejezéséhez, és hány szessziót vitt végig. |
| **Prioritás** | Kevésbé fontos |
| **Kiváltás** | 1. Sikeres bejelentkezés során |
| **Funkcionális követelmény** | A felhasználó láthatja számszerűen és egy fejlődésmutatón, hogy eddig hány adatot írt be. Minden sikeres leküldéssel növekszik ennek az értéke. Mikor a szessziót a felhasználó befejezte, egy üzenet jelenik meg, ami közli a sikeres kitöltést. |
| **Megszorítás** | 1. A böngészőben be van kapcsolva a JavaScript. |

5.6. Táblázat - Nyelvi funkció

|  |  |
| --- | --- |
| **Nyelvi Funkció** | |
| **Leírás** | Angol és magyar nyelvű felület biztosítása. |
| **Prioritás** | Kevésbé fontos |
| **Kiváltás** | 1. Az oldal bal felső sarkában lévő legördülő listából kiválaszt a felhasználó egy nyelvet. |
| **Funkcionális követelmény** | A nyelv váltás hatására az adott oldalon, és minden lapon, ahova az aktuális oldalról navigálunk, a nyelv megváltozik aszerint, hogy melyiket választotta a felhasználó. |
| **Megszorítás** | - |

5.7. Táblázat - Segítői funkció

|  |  |
| --- | --- |
| **Segítői Funkció** | |
| **Leírás** | A választott nyelven visszaküldött üzenet érvénytelen adatok esetén.  Leírása a kísérletnek, a letiltott és ignorált karaktereknek, valamint az egyéb lehetséges félreértések tisztázása, mind például a tanító jelszó és a bejelentkezési jelszó közti különbség.  Megjelenítése tanító módban a beírandó adatoknak. |
| **Prioritás** | Közepesen fontos |
| **Kiváltás** | 1. A felhasználó a regisztrációs vagy tanító oldalra navigál. |
| **Funkcionális követelmény** | Útmutatót adni arról, hogy miket kell tegyen a felhasználó, mik a lehetőségei és megszorításai. |
| **Megszorítás** | 1. A böngészőben be van kapcsolva a JavaScript. |

A felhasználói eset diagram itt látható: [diagram]

### Nem funkcionális követelmények

#### Skálázhatóság és továbbfejleszthetőség

Az alkalmazás legyen skálázható, és könnyen továbbfejleszthető. Ajánlott az MVC tervezési mintát alkalmazni.

#### Karbantarthatóság

A bekövetkezett hibák legyenek naplózva a könnyű karbantartás végett.

#### Nyelvi követelmények

Az oldal több nyelven is megírható, de minimálisan angol nyelven legyen megírva. Ajánlott a Laravel lokalizációs módszerét használva nyelvi fájlokban tárolni a megírt szöveget, a későbbi többnyelvűsítés érdekében.

### Interfész követelmények

#### Felhasználói interfész

A felhasználói interfész HTML-ben készüljön, böngészőből lehessen elérni. Ajánlott a HTML5 dokumentum típus, valamint Bootstrap 4.0 vagy annál újabb verzió, hogy könnyedén ki lehessen alakítani egy esztétikus felhasználói felületet. Emellett a Bootstrap segítségével könnyedén tudunk egy olyan nézetet létre hozni, melynek formája és elrendezése a használt eszköz kijelzőjétől függően változik

A felület vázlata a következőképpen néz ki: [ábra]

# Részletes tervezés

## Adatok tárolása Firebaseben

A felhasználói és kísérleti adatok tárolására a Firebase valós idejű NoSQL adatbázisára esett a választásunk. Ennek fő okai közé sorolható a nagyobb mennyiségű adatok gyors kezelése valamint a tulajdonsága, hogy a PHP asszociatív tömbjeit képes egyben feltölteni és átalakítani őket JSON formátumba, ez által nincs szűkség külön konverzióra.

A Firebase adatbázisban eltárolt csomópontoknak a következő a struktúrája: [ábra]

Magyarázat: Az egész adatbázis egy csomópont alatt helyezkedik el, ami a [*state-exam-project-keystroke*](https://console.firebase.google.com/u/0/project/state-exam-project-keystroke/database/state-exam-project-keystroke/data/). Ez alatt a következő csomópontok helyezkednek el:

1. *sessions* – a felügyelt tanításokhoz tartozó szesszió kódokat itt tároljuk
2. *user\_keystroke\_raw* – itt találhatóak a felhasználók nyers billentyűzési adatai, a következőképpen:
   1. *username(s)* – a felhasználónév alapján lehet megtalálni valakinek az adatait
      1. *timestamp(s)* – egy két részből álló azonosítója egy adott beküldésnek, az első része a beírás kezdeti dátumát, míg a második az aktuális időbélyeget tartalmazza
         1. *input\_field(s)* – az adott beküldés során kapott mezőket számokkal jelöltük, ahol 0 a felhasználónév, 1 a teljes név és 2 a tanítandó jelszó
            1. *keystroke(s)* – egy bemeneti mezőben leütött billentyű, aminek két értéke van: egy leütési és egy felengedési időbélyeg
      2. *forgeries* – a felhasználón végzett hamisítások száma
         1. *username(s)* – a hamisító vagy hamisítók neve, ez alatt ugyanolyan időbélyegek vannak, mint a 2.1. alatt
3. *user\_templates* – a felhasználók mintái, különböző algoritmusok által
   1. *username(s)* - a felhasználónév alapján lehet megtalálni valakinek a mintáit
      1. *algorithm(s)* – az algoritmus neve, amelyik által készült a minta
         1. *pattern* – a minta, ami általában pár vektorból áll
         2. *threshold* – a küszöbérték az adott felhasználóra
4. *users* – a felhasználók adatai, amiket a regisztrációnál adtak meg
   1. *username(s)* - a felhasználónév alapján lehet megtalálni valakit
      1. *device* – a felhasználó által választott eszköz
      2. *email* – a felhasználó e-mail címe
      3. *forgeries* – a hamisítások száma
         1. *username(s)* – a hamisítók nevei
            1. *number\_of\_samples* – az aktuális hamisító által leadott minták száma
      4. *full\_name* – a felhasználó teljes neve
      5. *input\_lengths* – a beírt karakterek hossza, mezőnként
         1. *input* – a mező számmal jelölve, mint a 2.1.1.1. csomópontnál, az adott mezőbe beírt karakterek hosszát tartalmazza
      6. *number\_of\_samples* – a felhasználó által beadott minták száma
      7. *password\_hash* – a bejelentkezéshez használt jelszó titkosítva
      8. *training\_password* – a tanuló jelszó, amit a felhasználó ír be több alkalommal
      9. *user\_age* – a felhasználó életkora
      10. *user\_gender* – a felhasználó neme
      11. *user\_nationality* – a felhasználó nemzetisége
      12. *username* – a felhasználónév

## Architektúra

## A rendszer tervezése és bemutatása

### A kontroller rész

### A modell rész

### A nézet (view) rész

### Segítő osztályok

## AJAX hívások bemutatása

## Adatok mozgása

# Gyűjtő üzembe helyezése és kísérleti eredmények

## Gyűjtő üzembe helyezése

## Felmerült problémák és megoldásaik

## Begyűjtött adathalmaz

### Felhasználóktól gyűjtött adathalmaz

### Imposztor adathalmaz

## Kísérleti eredmények

# Következtetések

## Megvalósítások

## Összehasonlítás hasonló rendszerekkel

## További fejlesztési irányok

# Irodalomjegyzék

[1] Morales, A., Falanga, M., Fierrez, J., Sansone, C., & Ortega-Garcia, J. (2015, September). Keystroke dynamics recognition based on personal data: A comparative experimental evaluation implementing reproducible research. In *Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on* (pp. 1-6). IEEE.

[2] Killourhy, K. S., & Maxion, R. A. (2009, June). Comparing anomaly-detection algorithms for keystroke dynamics. In *Dependable Systems & Networks, 2009. DSN'09. IEEE/IFIP international conference on* (pp. 125-134). IEEE.

[3] Teh, P. S., Teoh, A. B. J., & Yue, S. (2013). A survey of keystroke dynamics biometrics. The Scientific World Journal, 2013.

[4] Araújo, L. C., Sucupira, L. H., Lizarraga, M. G., Ling, L. L., & Yabu-Uti, J. B. T. (2005). User authentication through typing biometrics features. *IEEE transactions on signal processing*, *53*(2), 851-855.

# Függelékek