

**Politechnika Poznańska**  
**Wydział Budowy Maszyn i Zarządzania**  
**Katedra Zarządzania i Inżynierii Produkcji**



PRACA DYPLOMOWA MAGISTERSKA

**Budowa urządzenia komunikacji głosowej  
w warunkach produkcyjnych dla osób  
głuchoniemych.**

**Paweł CZEKAŁA**  
**Krzysztof PASZAK**

**Kierunek:** Zarządzanie i Inżynieria Produkcji

**Specjalność:** Informatyzacja Produkcji

Promotor:  
**dr inż. Filip GÓRSKI**

Poznań 2019

MIEJSCE NA KARTĘ PRACY DYPLOMOWEJ

# Spis treści

Streszczenie .....	4
Summary .....	4
Podział pracy .....	5
1. Wprowadzenie.....	6
2. Cel pracy i zakres pracy .....	8
3. Analiza literatury.....	9
3.1. Osoby niesłyszące na rynku pracy.....	9
3.2. Język migowy w przemyśle.....	10
3.2.1. Wprowadzenie.....	10
3.2.2. Podstawowe definicje.....	12
3.2.3. Funkcjonowanie osób głuchoniemych w przemyśle.....	14
3.3. Komunikacja bezprzewodowa.....	16
3.4. Klasyfikacja gestów.....	20
3.4.1. Wprowadzenie.....	22
3.4.2. Metody komunikacji wspomagającej i alternatywnej .....	20
3.4.3. Klasyfikatory .....	22
3.4.4. Metody klasyfikacji.....	24
3.4.5. Strategie rozwiązywania problemów wieloklasowych .....	27
3.5. Analiza istniejących rozwiązań .....	28
3.5.1. Wprowadzenie.....	28
3.5.2. Systemy wizyjne .....	29
3.5.3. Urządzenia wearable .....	31
3.5.4. Porównanie rozwiązań .....	33
4. Metodyka pracy.....	35
4.1. Koncepcja i plan prac .....	35
4.2. Struktura rękawicy oraz oprogramowania.....	38
4.2.1. Zastosowane komponenty sprzętowe.....	38
4.2.2. Schemat elektryczny .....	40
4.2.3. Software .....	41
4.3. Budowa oprogramowania sterującego.....	43
4.3.1. Struktura gestu.....	43
4.3.2. Baza słów środowiska produkcyjnego .....	46
4.4. Metodyka testowania rozwiązania.....	47
5. Wyniki pracy .....	51
5.1. Uzyskany prototyp rozwiązania .....	51
5.2. Testy i ocena klasyfikacji .....	55
5.3. Wyniki ankiety i testów grupy docelowej .....	59
5.4. Dyskusja i wnioski.....	66
6. Podsumowanie .....	70
Literatura .....	72
Spis Rysunków.....	75
Spis tabel .....	77
Załącznik 1. Schemat elektryczny urządzenia .....	78
Załącznik 2. Ankieta .....	79
Załącznik 3. Spis zawartości płyty CD dołączonej do pracy .....	85

## Streszczenie

Temat pracy dotyczy jednego z kluczowych aspektów funkcjonowania jednostki w grupie, a więc komunikacji, bez której życie zarówno zawodowe jak i osobiste narażone byłoby na szereg nieporozumień i komplikacji. Dla osób z niepełnosprawnościami słuchowymi proces komunikowania się w przedsiębiorstwie produkcyjnym z osobami słyszącymi jest szczególnie utrudniony, dlatego podjęto próbę stworzenia prototypu rękawicy współpracującej z aplikacją mobilną, której zadaniem będzie tłumaczenie języka migowego na tekst i mowę.

W początkowych dwóch rozdziałach opisano cel oraz główne założenia, które zostały postawione przez autorów pracy oraz które zostały skonsultowane z promotorem. W kolejnym rozdziale skupiono się na analizie literatury pod kątem specyfiki posługiwania się językiem migowym w środowisku produkcyjnym oraz na możliwościach technologicznych wykorzystanych w istniejących rozwiązaniach a także w niniejszej pracy. W dalszej części przedstawiona została metodyka pracy począwszy od koncepcji rozwiązania aż po sposoby przeprowadzania testów na gotowym urządzeniu. Ostatnie dwa rozdziały prezentują i podsumowują efekty każdego etapu, który został zrealizowany w trakcie tego projektu.

## Summary

The subject of the masters' thesis deals with one of key aspects of functioning of the unit in a group, communication, without which both professional and personal life would be exposed to a number of misunderstandings and complications. For people with auditory disabilities, a process of communicating with hearing people in a manufacturing company is especially difficult, therefore an attempt of creating a prototype of a glove cooperating with a mobile application, which task will be to translate sign language into text and speech was made.

In first two chapters the main target and assumptions, that were posed by the authors and consulted with the promoter, are described. In next chapter focus was placed at analysis of literature, especially considering specifics of using sign language in manufacturing environment and technological possibilities used in this project as well as in existing solutions. In the following part of this thesis the methodology was featured, starting with conception of solution and ending with ways of carrying out tests on finished device. Last two chapters present and sum up effects of every stage that was made during this project.

# Podział pracy

Paweł Czekala:

- Analiza istniejących sensorów i rozwiązań elektronicznych,
- Budowa prototypu urządzenia – część elektroniczna,
- Opracowanie technologii zbierania i obróbki danych z sensorów,
- Opracowanie algorytmu klasyfikacji gestów i bazy słów,
- Badanie i ocena skuteczności rozwiązania,
- Autorstwo rozdziałów: 2, 3.1, 3.4, 4.1, 4.3.1, 4.4, 5.2, 5.3, 5.4, 6.

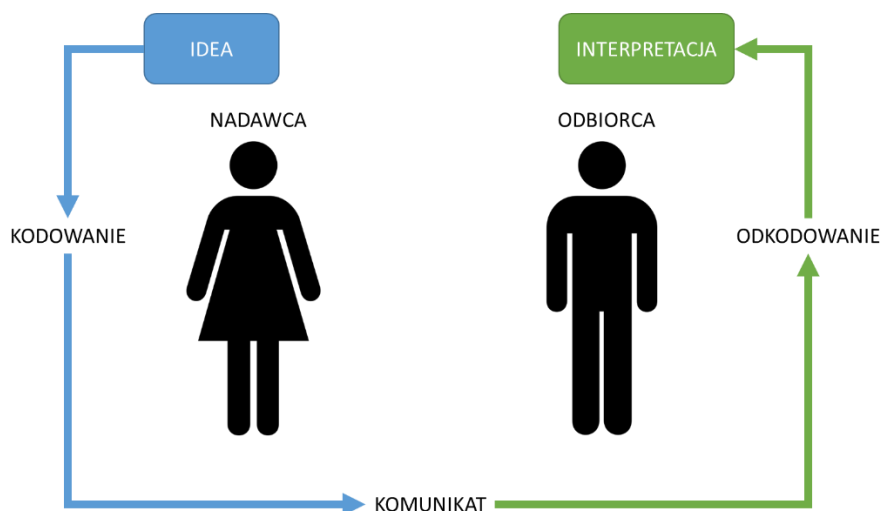
Krzysztof Paszak:

- Analiza istniejących rozwiązań konstrukcyjnych, dobór komponentów prototypu,
- Analiza istniejących sensorów i rozwiązań elektronicznych,
- Budowa prototypu urządzenia – część mechaniczna,
- Budowa aplikacji mobilnej dla użytkownika, sterującej działaniem urządzenia,
- Autorstwo rozdziałów: 1, 3.2, 3.3, 3.5, 4.2, 4.3.2, 5.1, 5.4, 6.

# 1. Wprowadzenie

W środowisku pracy można spotkać wiele różnych osób, ale najczęściej są to ludzie słyszący. Czasami jednak zdarza się, że nawet z nimi trudno jest się porozumieć, mimo tak oczywistego kanału komunikacyjnego jakim jest mowa. Odbiór informacji głosowej jest odruchem bezwarunkowym, ale nie świadczy o pomyślnej komunikacji między ludźmi. Najczęściej osobowość drugiego człowieka, jego sposób bycia oraz nastawienie decydują, że dana osoba jest warta kontaktu. Ale jak poznać charakter danej osoby nie rozmawiając z nią? Odpowiedź na to pytanie jest kluczowa dla osób niesłyszących, ponieważ mowa jest dla nich swoistym językiem obcym, którego nie są w stanie się nauczyć. Znalezienie odpowiedniej metody komunikacji jest zadaniem trudnym, ale możliwym do realizacji. W ciągu ostatnich lat zostały wypracowane techniki i stworzone rozwiązania, w znacznym stopniu pomagające osobom słyszącym zrozumieć głuchych.

W obliczu stale rosnącej liczby osób niesłyszących i niedosłyszących w otaczającym społeczeństwie, niewątpliwie coraz bardziej istotną kwestią funkcjonowania w życiu zawodowym tej grupy ludzi stanowić będzie aspekt komunikacji międzyludzkiej. Na rysunku 1.1 przedstawiony został schemat przekazywania komunikatu. Widać na nim, że do odebrania przez odbiorcę informacji od nadawcy potrzebne jest zarówno kodowanie jak i odkodowanie. W sytuacji osób z niepełnosprawnościami słuchowymi, w wielu przypadkach wiąże się to z upośledzeniem ścieżki wymiany komunikatów zarówno wysyłanych jak i odbieranych przez taką osobę. Oznacza to, że wysyłając komunikat w postaci gestów dłoni, ruchów ciała i mimiki twarzy taka osoba jest co prawda w stanie przekazywać informację osobie słyszącej (nieznającej języka migowego), ale brak możliwości prawidłowego odkodowania takiej wiadomości powoduje u odbiorcy efekt niemal całkowitego zaszumienia komunikatu. W takiej sytuacji można mówić tylko o przekazywaniu podstawowych i oczywistych informacji. W tej samej grupie ludzi występuje problem także przy próbie przekazania komunikatu w drugą stronę: od osoby słyszącej do niesłyszącej. Tutaj również istnieje możliwość przekazywania jedynie prostych komunikatów. Wyjątkiem może być jednak umiejętność czytania z ruchu warg przez osobę niesłyszącą, lecz w wielu przypadkach (jak chociażby jednoczesna komunikacja z większą liczbą osób) również taka zdolność nie pozwoli na płynne przekazywanie komunikatów.



Rys. 1.1. Przepływ komunikatu między nadawcą a odbiorcą (opracowanie własne na podstawie (Płuska, 2017))

Głównym powodem stworzenia projektu urządzenia komunikacji głosowej była potrzeba znalezienia odpowiedniego kanału wymiany informacji pomiędzy słyszącymi i niesłyszącymi w zakładach produkcyjnych. To środowisko jest szczególne, ponieważ stawia przed urządzeniami duże wymagania. Mnogość warunków zakłócających, występująca na halach produkcyjnych, nie pozwala na wykorzystanie wielu rozwiązań technicznych. Postanowiono zaprojektować oraz zbudować urządzenie komunikacji głosowej, ułatwiające niesłyszącym oraz niedosłyszącym porozumiewanie się z osobami słyszącymi, które nie znają gestów języków migowych.

## 2. Cel pracy i zakres pracy

Głównym celem niniejszej pracy dyplomowej było wykonanie urządzenia komunikacji głosowej dla osób głuchoniemych w warunkach produkcyjnych. Do celów szczegółowych należy zaliczyć:

- dobór komponentów i parametrów sprzętowych,
- opracowanie algorytmów rozpoznających gesty użytkownika na wysokim poziomie skuteczności,
- stworzenie bazy słów charakterystycznych dla środowiska produkcyjnego,
- opracowanie i zaprogramowanie aplikacji mobilnej,
- budowę rękawic, integrację warstwy sprzętowej i programowej.

Zakres pracy obejmował:

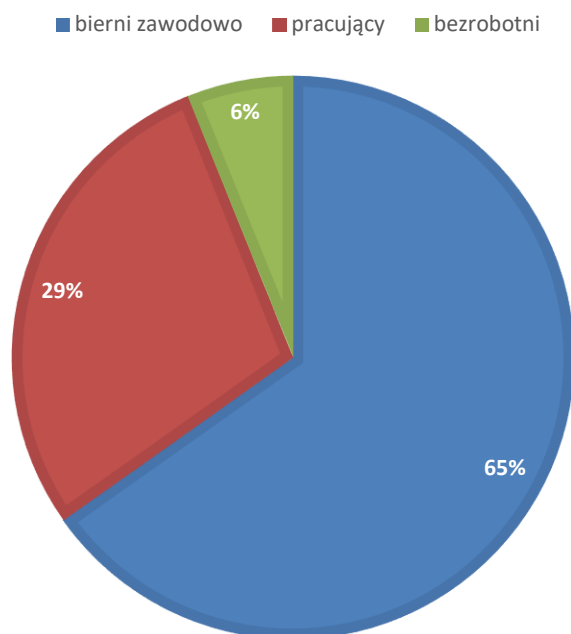
- analizę literaturową z zakresu funkcjonowania języka migowego w przemyśle, komunikacji bezprzewodowej, klasyfikacji gestów oraz analizy istniejących rozwiązań,
- projekt i konstrukcję urządzenia,
- opracowanie algorytmów klasyfikacyjnych,
- wykonanie interfejsu użytkownika w postaci aplikacji mobilnej,
- przeprowadzenie ankiety, której wyniki pozwolą m.in. wskazać mocne i słabe punkty wykonanego prototypu.



## 3. Analiza literatury

### 3.1. Osoby niesłyszące na rynku pracy

Celem niniejszej pracy było opracowanie takiego kanału komunikacyjnego, który umożliwiłby osobom z niepełnosprawnościami słuchowymi zarówno odbierać jak i przekazywać komunikaty w kontaktach z osobami słyszącymi, nieznającymi języka migowego. Wymiana komunikatów takim kanałem umożliwiłaby, osobom niesłyszącym i niedosłyszącym w środowisku produkcyjnym, łatwiej przekazywać swoje potrzeby, a także w większym stopniu rozumieć i odbierać wysyłane do nich informacje. Na podstawie danych GUS, zobrazowanych na rysunku 3.1, widać jak liczną procentowo grupę stanowią osoby z niepełnosprawnością słuchową, które są bierne zawodowo: 65,2%. Podstawową przyczyną tak dużego odsetku tej grupy ludzi są czynniki wynikające z barier komunikacyjnych. Proces tworzenia i rozumienia mowy przez osobę głuchą potrafi być dużym obciążeniem zarówno fizycznym, jak i psychicznym. Dodatkowym utrudnieniem jest fakt, iż gramatyka języka migowego jest odmienna od gramatyki języka narodowego, można zatem przyjąć, że dla osoby niesłyszącej język narodowy, powszechnie stosowany w miejscu pracy jest językiem obcym, tak jak, np. dla Polaka język angielski (Sak, 2012). Gdyby udało się przełamać barierę komunikacyjną, wówczas zwiększyłyby się szanse na zmniejszenie bariery izolacji życia zawodowego osób niesłyszących i niedosłyszących.



Rys. 3.1. Rozkład osób z niepełnosprawnością słuchową na rynku pracy (opracowanie własne na podstawie (GUS, 2012))

Opracowany kanał informacyjny powinien przyjąć formę urządzenia, które poza funkcjami komunikacyjnymi, dodatkowo nie utrudniałoby pracy w środowisku produkcyjnym. W takim przypadku urządzenie nie może być stacjonarne, tylko możliwie jak najbardziej mobilne, aby zapewnić użytkownikowi swobodę wykonywanych w pracy ruchów. Środowisko produkcyjne charakteryzuje się obecnością maszyn i robotów oraz zintensyfikowanego przepływu różnego rodzaju zasobów i surowców, a co za tym idzie, konstruowane urządzenie powinno spełniać również wszelkie aspekty bezpieczeństwa i nie utrudniać użytkownikowi przestrzegania przepisów BHP.

## **3.2. Język migowy w przemyśle**

### **3.2.1. Wprowadzenie**

Komunikacja osób słabosłyszących oraz głuchych w środowiskach produkcyjnych jest trudnym wyzwaniem. Duża liczba osób, wielkie przestrzenie oraz wiele źródeł dźwięku utrudniają odbiór komunikatów. Na przestrzeni lat stworzono kilka metod usprawniających komunikację osób z niepełnosprawnościami słuchowymi w miejscach pracy. Do takich rozwiązań można zaliczyć:

1. Telefony ze wzmocnioną głośnością i cewką indukcyjną. Jest to rozwiązanie, które należy do środków technicznych i służy wyłącznie osobom słabosłyszącym. Urządzenie wspomaga komunikację „jeden z jednym”. Główną funkcją zastosowania pętli indukcyjnej jest wyeliminowanie zakłóceń i szumów w rejestrowanym dźwięku.
2. Urządzenie wspomagające słyszenie. Poprawia transmisję głosu, który przekazywany jest radiowo między dwoma urządzeniami. Nadawca nosi urządzenie FM (ang. *Frequency Modulation*) wyposażone w nadajnik i mikrofon, natomiast odbiorca posiada odbiornik sygnału FM oraz wzmacniacz. Odebrany i wzmocniony sygnał audio jest przekierowany do wbudowanego w urządzeniach głośnika. Taka metoda komunikacji nie sprawdza się jednak w warunkach, w których istnieje wiele źródeł zakłóceń dźwiękowych oraz wiele interferujących pól magnetycznych i radiowych.
3. Stenotypia. Jest to technika, w której konwersacja zapisywana jest w czasie rzeczywistym za pomocą maszyny stenograficznej (lub komputera). Zapis może być wyświetlany za pomocą projektora, a osoba niesłysząca jest w stanie kontrolować rozmowę z niewielkim opóźnieniem związanym z szybkością rejestrowania rozmowy przez stenotypistę na klawiaturze. Dużą zaletą takiego rozwiązania jest wierne

przekazywanie niezmienionej treści rozmowy osobie niesłyszącej. Słabymi stronami stenotypii jest stosunkowo wysoki koszt potrzebnych zasobów, zarówno rzeczowych (maszyna stenograficzna, komputer, projektor) jak i osobowych (Sak, 2012, str. 33-34).. Do współczesnych wersji stenotypii można zaliczyć, np. asystenta Google (dla Android) czy Siri (dla Apple). Przy użyciu sprzętu wyposażonego w mikrofon, najczęściej telefonu komórkowego, oprogramowania te są w stanie przełożyć wypowiedź ustną na tekst, ukazujący się na ekranie. Szczególnym urządzeniem, które wykorzystuje rozpoznawanie mowy w czasie rzeczywistym jest LTCCS (ang. *Live Time Closed Captioning System*). Jest to urządzenie, które za pomocą mikrofonu i systemu procesorowego rozpoznaje mowę i wyświetla ją na okularach rzeczywistości rozszerzonej (Coxworth).

4. Tworzenie skróconych notatek. Jest to uproszczona wersja stenotypii. Pracownik tworzy na papierze, komputerze lub telefonie skrócone notatki z rozmowy i przekazuje je osobie niesłyszącej. Zaletą tego rozwiązania jest niski koszt i łatwa dostępność sprzętu (telefon, papier), jednak wadą jest utrata części informacji zawartych w pełnej konwersacji.
5. Komunikacja drogą elektroniczną. Opiera się na przesyłaniu komunikatów przez SMS, e-mail lub komunikator internetowy. Metody te są coraz bardziej popularne i łatwo dostępne, jednak opierają się tylko na kontakcie wirtualnym.
6. Tłumaczenia języka migowego. Jest to najbardziej pożądana przez osoby głuche metoda komunikacji. Tłumaczenia z języka polskiego na język migowy i *vice versa* są możliwe poprzez zatrudnieniu tłumacza. Dzięki temu niesłyszący pracownik jest w stanie w pełni uczestniczyć w życiu przedsiębiorstwa. Takie rozwiązanie generuje jednak wysokie koszty związane z utrzymaniem tłumacza języka migowego.
7. Tłumaczenia online. Jest to coraz popularniejsza metoda komunikacji osób z niepełnosprawnościami słuchowymi, które wykorzystując odpowiednią aplikację, zainstalowaną na komputerze lub telefonie komórkowym, są w stanie porozumiewać się z osobami słyszącymi, które nie znają języka migowego. Tłumaczenie jest realizowane za pomocą wideokonferencji z tłumaczem języka migowego. Takie rozwiązanie jest trudne do zaimplementowania w przedsiębiorstwie produkcyjnym, ale w odpowiednich warunkach możliwe do zrealizowania, np. w przypadku kooperacji zespołów projektowych (Sak, 2012, str. 33-34).

Zdobyte techniki pozwalają ciągle ulepszać i odkrywać nowe technologie wspomagające narząd słuchu osób słabosłyszących, natomiast dla ludzi głuchych najistotniejszy jest dostęp do

tłumaczeń z języka migowego (PJM) na język narodowy i *vice versa*. Spośród wymienionych wyżej metod, w warunkach produkcyjnych można pomyślnie wykorzystywać telefony ze wzmocnioną głośnością oraz cewką eliminującą zakłócenia dźwiękowe, skrócone notatki na papierze lub telefonie komórkowym oraz tłumaczenia z pracownikiem znającym język migowy. Pozostałe rozwiązania nie znajdują wykorzystania w zakładach produkcyjnych, ponieważ generują duże koszty lub ich warunki techniczne sprawiają, że nie są odporne na liczne źródła zakłóceń (Sak, 2012, str. 35).

### **3.2.2. Podstawowe definicje**

Stopień, w jakim uszkodzony jest słuch człowieka, definiuje jego przynależność do danej grupy. Poniżej przedstawiono definicje klasyfikujące ludzi z niepełnosprawnością słuchową.

1. Osoba głucha. W ten sposób określa się osobę, która mimo stosowania aparatu słuchowego nie jest w stanie zrozumieć mowy ludzkiej. Dla takiego człowieka słuch nie posiada znaczenia praktycznego (Gałkowski i in., 1978, str. 24).
2. Osoba głuchoniema. Jest to człowiek, który na skutek głuchoty nie opanował umiejętności mowy, i co za tym wynika – nie posługuje się nią, ani nie potrafi odczytywać słów z ruchu ust (Gałkowski i in., 1978, str. 24).
3. Osoba niedosłysząca lub inaczej słabosłysząca. Takim terminem nazywa się osobę z uszkodzonym słuchem w stopniu ograniczającym odbiór mowy drogą słuchową. Taki człowiek ma szansę opanować mowę drogą naturalną (Gałkowski i in., 1978, str. 24).
4. Osoba niema. Jest to osoba słysząca, która nie opanowała aparatu mowy (Gałkowski i in., 1978, str. 25).
5. Osoba niesłysząca. Nazywa się tak każdą osobę z uszkodzeniem słuchu w stopniu znacznym i głębokim (Gałkowski i in., 1978, str. 24).

Klasyfikacja niepełnosprawności narządu słuchu według Międzynarodowego Biura Audiofonologii (BIAP) wyróżnia cztery stopnie uszkodzenia słuchu:

1. Uszkodzenie słuchu w stopniu lekkim. Wiąże się z ubytkiem od 20dB do 40dB. Osoby z taką wadą mogą mieć utrudnione słyszenie w hałasie oraz dużych przestrzeniach i na dużych odległościach. Wskazaniem dla nich jest używanie aparatów słuchowych.
2. Uszkodzenie słuchu w stopniu umiarkowanym. Jest to ubytek pomiędzy 40db a 70dB. Osoby z taką wadą nazywa się niedosłyszącymi lub słabosłyszącymi, często korzystają z aparatów słuchowych oraz innych pomocy technicznych. Do prawidłowego słyszenia potrzebują bardzo dobrych warunków akustycznych.

3. Uszkodzenie w stopniu znacznym. Charakteryzuje się ono ubytkiem słuchu w granicach od 70dB do 90dB. Ludzie z uszkodzeniem słuchu w stopniu znacznym nawet z wykorzystaniem aparatów słuchowych mają problem ze zrozumieniem mowy ludzkiej.
4. Uszkodzenie słuchu w stopniu głębokim. Wiąże się z ubytkiem powyżej 90dB i brakiem możliwości rozpoznania mowy nawet przy wykorzystaniu jakiegokolwiek pomocy technicznej wspomagającej słuch (Kozicka, 2011).

PJM to Polski Język Migowy. „Ustawa z dnia 19 sierpnia 2011r. o języku migowym i innych środkach komunikowania się (Dz. U. z 2011r. Nr 2009, poz. 1243) definiuje polski język migowy jako naturalny, wizualno-przestrzenny język komunikowania się (Dunaj, 2015).

PJM powstawał sukcesywnie z biegiem lat, a stworzony system znaków, łączył w sobie dwa elementy językowe: gramatykę i słownictwo. Dziś liczba tworzonych struktur językowych jest nieograniczona. W PJM główną formą przekazywania słów jest odpowiedni ruch, ułożenie palców, lokalizacja ręki oraz mimika. Odmienność tej formy komunikacji powodowała, że była ona traktowana jako sposób porozumiewania się osób głuchych a nie jako język. Jednak po wielu badaniach i publikacjach na temat gramatyki PJM, zmienił się sposób postrzegania tej formy komunikacji (Dunaj, 2015).

SJM to System Językowo Migowy. Został stworzony w celu ułatwienia edukacji głuchych, a w szczególności nauki języka polskiego, dlatego SJM opiera się całkowicie na gramatyce języka polskiego. Tekst przekazywany za pomocą SJM nie musi być tłumaczony na język polski, ponieważ zdania w języku polskim i SJM brzmią dokładnie tak samo. Ruchy rąk (gesty) używane są równocześnie z mową. Jedno słowo ma swoje konkretne znaczenie, w zależności od kontekstu, a szyk zdania jest zgodny z gramatyką języka polskiego. SJM jest sukcesywnie wypierany przez PJM, dlatego coraz mniej osób niesłyszących uczy się tego system znaków. (Grabowska-Orządzała i in., 2012, str. 41-42).

Kulturowy model głuchoty to koncepcja, która nie uznaje głuchoty jako niepełnosprawności, a także odrzuca możliwości jej leczenia. Według tej myśli głuchota to rodzaj tożsamości, autonomii. Każdy, kto ma problem z komunikacją słuchowo-werbalną należy do społeczności głuchych, która ma własny język, wartości oraz specyficzne zasady zachowania. Fundamentem kulturowego modelu głuchoty jest język migowy, który przekazywany jest z pokolenia na pokolenie i jest nośnikiem dziedzictwa głuchych (Sak, 2012, str. 15-16).

### 3.2.3. Funkcjonowanie osób głuchoniemych w przemyśle

Sytuacja osób z niepełnosprawnością słuchową na rynku pracy a w szczególności w zakładach produkcyjnych zależy od wielu elementów, do których zaliczyć można m.in.:

- nastawienie,
- zdrowie,
- edukację,
- sytuację ekonomiczną kraju,
- stereotypowe myślenie na temat głuchych przez osoby słyszące.

Według głuchych to właśnie ostatni czynnik ma najważniejsze znaczenie przy próbie nawiązania współpracy na linii pracodawca – osoba głucha. Gdyby problem komunikacji był główną barierą w podjęciu pracy, to zapewnienie tłumaczenia powinno rozwiązać ten kłopot. Jednak największą barierą jest brak wystarczającej wiedzy na temat tego typu niepełnosprawności. Już samo tytułowanie takich osób niepełnosprawnymi konotuje wiele negatywnych odczuć. Osoby głuche są gotowe do podjęcia pracy i co najważniejsze posiadają umiejętności do wykonywania powierzonych im zadań (Butkiewicz i in., 2014a, str. 89-90). Niesłyszący posiadają bardzo cenne dla pracodawców umiejętności:

- skoncentrowanie się na zadaniu,
- doskonała obserwacja, analiza otoczenia,
- motywowanie się do wykonywania pracy (szczególnie u osób młodych),
- dobra pamięć wzrokowa.

Jednak posiadają również pewne ograniczenia:

- w komunikacji,
- w myśleniu abstrakcyjnym,
- w rozumieniu złożonych tekstów pisanych.

W przedsiębiorstwach produkcyjnych niewskazane jest podejmowanie przez głuchych pracy w hałasie (szczególnie dźwięki o małej częstotliwości), a także w przestrzeniach o zwiększonym lub obniżonym ciśnieniu atmosferycznym (Butkiewicz i in., 2014b, str. 19-21).

Procedura zatrudnienia pracownika z niepełnosprawnością słuchową na stanowisko produkcyjne jest identyczna jak dla osoby pełnosprawnej. Pracodawca przeprowadza rozmowę kwalifikacyjną z kandydatem. Po podjęciu decyzji o przyjęciu osoby głuchej wystawia skierowanie do lekarza medycyny pracy, które zawiera dokładny opis stanowiska, na które zostaje przyjęty, wraz z oceną ryzyka sporządzoną przez dział BHP. Lekarz na podstawie skierowania ocenia aktualny stan kandydata a także próbuje ustalić, czy istnieje ryzyko

pogorszenia się stanu zdrowia w przypadku zatrudnienia na danym stanowisku. Osoba głucha na drodze do zatrudnienia musi przekonać do swojej osoby i posiadanych umiejętności pracodawcę, specjalistę BHP oraz lekarza medycyny pracy. Według głuchych, specjaliści skupiają się tylko i wyłącznie na braku słuchu i konsekwencjach z tym związanych, dlatego wielu z nich nie otrzymuje pozytywnej opinii lekarzy (Butkiewicz i in., 2014b, str. 25-28).

Prace zakwalifikowane przez Ministerstwo Zdrowia jako wymagające odpowiedniej sprawności narządu słuchu to m.in.:

- operowanie wózkiem jezdniowym,
- operowanie urządzeniami podnośnikowymi,
- obsługa maszyn w ruchu,
- praca na wysokości.

Zabroniona obsługa maszyn w ruchu jest dużym ograniczeniem dla osób głuchych starających się o pracę w przedsiębiorstwach produkcyjnych. Często nie otrzymują oni zdolności do wykonywania pracy właśnie ze względu na ten czynnik. Medyczny punkt widzenia skupia się na przeszkodach związanych z głuchotą, pomijając możliwości osób głuchych, które często posiadają silniej wykształcone pozostałe zmysły. Nie tylko wzrok, ale również dotyk pozwala osobie niesłyszącej rozpoznać zagrożenia wynikające, np. ze zmiany stanu pracy maszyny na stanowisku produkcyjnym. Aby opinia lekarza medycyny pracy była pozytywna, kluczowym jest szczegółowe opisanie w skierowaniu na badania lekarskie sposobu przystosowania stanowiska pracy dla osoby niesłyszącej (Butkiewicz i in., 2014b, str. 28-30).

Liczba osób głuchych pracujących w przedsiębiorstwach produkcyjnych jest niewielka i wynosi zaledwie 18%. Głównymi przyczynami są:

- niewiedza dotycząca przepisów związanych z pobieraniem świadczeń rentowych (obawa utraty renty),
- konieczność ciągłego doksztalcania się,
- odpowiedzialność za wykonywaną pracę,
- zmianowy charakter pracy,
- konieczność kontaktowania się z klientami,
- odległość od miejsca zamieszkania do pracy,
- obawa przed byciem jedynym głuchym w miejscu pracy.

Realizacja projektów aktywizacji zawodowej osób głuchych wskazała język komunikacji jako istotny czynnik różniący szanse osób głuchych na rynku pracy. Dlatego ważnym staje się stworzenie rozwiązania, technologicznego lub organizacyjnego, które pomogłoby stworzyć

wspólny kanał komunikacyjny pomiędzy osobami słyszącymi i głuchymi (Butkiewicz i in., 2014a, str. 92 – 102).

### 3.3. Komunikacja bezprzewodowa

Urządzenie komunikacji głosowej, które jest tematem pracy zostało zaprojektowane w sposób, który pozwala całkowicie wyeliminować potrzebę używania przewodów zasilających oraz transmisyjnych. Oczywistym jest fakt, że komunikacja przewodowa jest o wiele bezpieczniejsza (dane nie zostaną przechwycone) i pewna (nie są tracone informacje), ale niesie ze sobą wiele ograniczeń i niedogodności. Mając powyższe na względzie, postanowiono oprzeć wymianę danych pomiędzy telefonem komórkowym a mobilną rękawicą na komunikacji bezprzewodowej. Oferowane na rynku moduły komunikacyjne dawały wybór kilku możliwości.

Istnieje wiele rozwiązań w dziedzinie bezprzewodowych połączeń, jednak należy przeanalizować parametry każdej technologii, aby wybrać najkorzystniejszą opcję, która w pełni zrealizuje założenia projektu. Kluczowymi parametrami są:

1. Przepustowość danych. Ma ona istotny wpływ, ponieważ określa ilość przesyłanych danych w jednostce czasu. Ważne jest, aby w założeniach projektu ustalić jaka szybkość transferu danych będzie wystarczająca dla danej aplikacji.
2. Zasięg. Określa on maksymalną możliwą odległość pomiędzy urządzeniami mobilnymi, przy której zachowane jest poprawne połączenie bezprzewodowe.
3. Pobór prądu. Jest to aktualnie jeden z najistotniejszych czynników, mający bezpośredni wpływ na maksymalny czas pracy urządzenia na baterii. Dąży się do wykorzystania energooszczędnych urządzeń i takich metod, które w jak najefektywniejszy sposób zarządzają pobieraną przez urządzenie energią.
4. Niezawodność transmisji. Jest ważnym czynnikiem wpływającym na poprawność przesyłanych danych. Niestabilna transmisja podatna na zakłócenia zewnętrzne może wprowadzić duże błędy, które będą fałszować otrzymane dane.

Na początku odrzucono trzy technologie, które posiadają wiele ograniczeń przez co są używane w wąskiej dziedzinie aplikacji: podczerwień (IR, ang. *infrared*), łączność komórkowa (GSM, ang. *Global System for Mobile Communications*) oraz komunikacja bliskiego zasięgu NFC (ang. *Near-field communication*).

Komunikacja przez podczerwień była w pewnym okresie popularna przy przesyłaniu niewielkich plików danych – najczęściej były to obrazy oraz pliki audio. Współczesne telefony komórkowe w znacznej większości nie posiadają nadajnika/odbiornika podczerwieni, głównie



z uwagi na bardzo niską przepustowość danych i wysoki zakres wymagań. Komunikacja przez podczerwień wymagała idealnego ustawienia odbiornika, w bardzo bliskiej odległości, w stosunku do nadajnika (Brigance, 2017).

Komunikacja urządzeń przez sieć komórkową jest rzadko używana w projektach z urządzeniami typu *wearables* (urządzenia ubieralne, komputer ubierany), ponieważ generowałyby dodatkowe koszty związane z opłatami abonamentowymi. Innymi wadami tej technologii są: duży pobór prądu modułów GSM oraz opóźnienia w transmisji danych (Brigance, 2017).

Technologia NFC w ostatnim czasie stała się bardzo popularna za sprawą możliwości płatności zbliżeniowych, jednak w aplikacjach użytkowych jej główną wadą jest bardzo krótki zasięg, do ok. 20mm (Brigance, 2017).

Bluetooth należy do technologii WPAN (ang. *Wireless personal area network*) i jest technologią bezprzewodową, gdzie transmisja danych odbywa się za pomocą fal radiowych UHF, o krótkiej długości w paśmie ISM od 2,4 do 2,485 GHz. Bluetooth został wynaleziony w 1994r. przez firmę Ericsson, działającą w branży urządzeń i usług sieciowych oraz telekomunikacyjnych. Obecnie zarządzany jest przez Bluetooth Special Interest Group (SIG). Była to bezprzewodowa alternatywa dla standardu komunikacji przewodowej RS-232 (Buenafior, 2019).

Bluetooth opiera się na technologii fal radiowych niskiej mocy, co powoduje, że osiąga krótki zasięg i średnią prędkość transmisji. Główną zaletą tej technologii jest niski pobór mocy, co oznacza, że może być wykorzystywany w urządzeniach zasilanych bateryjnie (Man, 2002, str. 3).

Od 1994 do 2016 roku powstało pięć wersji komunikacji bezprzewodowej Bluetooth, z których każda następna przynosiła kolejne ulepszenia parametrów komunikacji:

1. Bluetooth 1.x. Był pierwszą wersją, posiadającą wiele błędów ze zgodnością oraz charakteryzował się niską maksymalną przepustowością danych, tj. 1Mb/s. Urządzenia z tą wersją praktycznie już nie występują.
2. Bluetooth 2.x. Wydany został przez grupę Bluetooth Special Interest Group. Nowa wersja przyniosła poprawki, w wyniku których zwiększono szybkość transmisji danych do poziomu 3Mb/s. Była to jedna z najbardziej popularnych wersji bluetooth i używana jest do tej pory w wielu urządzeniach.
3. Bluetooth 3.x. Powstał w roku 2009 i posiadał dodatkowe oznaczenie HS co oznaczało *High Speed* (wysoka prędkość). Bluetooth w nowej wersji pozwalał na transfer danych do 24Mb/s, dzięki uruchomieniu alternatywnego radia, takiego jak

w urządzeniach wykorzystujących WiFi. Ta wersja wprowadziła wiele ulepszeń związanych z używaniem adresów w protokole, a także aktualizacje funkcji kontroli mocy anteny.

4. Bluetooth 4.x. Powstał w roku 2010 i był swego rodzaju przełomem, ponieważ wprowadził obsługę zbierania danych z urządzeń o niskim zużyciu energii (tzw. LE - *Low Energy*). Ta funkcja została zaprojektowana dla branży opieki zdrowotnej, fitnessu oraz rozrywki domowej. Praktycznie każde zakupione teraz urządzenie Bluetooth wykorzystuje BLE (ang. *Bluetooth Low Energy*) (Buenaflor, 2019). BLE zostało stworzone dla urządzeń, które muszą działać przez długi czas na małych źródłach energii. BLE znajduje zastosowanie w aplikacjach, w których otrzymywane są niewielkie aktualizacje danych, takie jak bieżące tętno, z odczytem co kilka sekund. Proces „parowania” jest uniwersalny, dzięki czemu urządzenia BLE można podłączyć do urządzenia centralnego, takiego jak smartfon (Brignace, 2017).
5. Bluetooth 5.x. Ogłoszony został w czerwcu 2016r. Zgodnie z zapewnieniem ma znacznie większy zasięg, prędkość i możliwości przesyłania większych paczek danych w jednostce czasu. Celem wprowadzenia Bluetooth v5, bo tak nazwano najnowszą wersję, było stworzenie niezawodnego i szybkiego połączenia pomiędzy urządzeniami mobilnymi (Buenaflor, 2019).

Kolejną konkurencyjną technologią bezprzewodową, którą często można spotkać w aplikacjach, jest WiFi (ang. *Wireless Fidelity*). Jest to standard opisujący szybkie połączenie bezprzewodowe na stosunkowo krótkich dystansach pomiędzy urządzeniami takimi jak laptopy, telefony komórkowe, routery. WiFi to typ bezprzewodowej sieci, często stosowany do komunikacji między elektroniką użytkową (np. telewizory, odtwarzacze DVD, aparaty, telefonu komórkowe) (Agarwal, 2018).

Sieć WiFi wykorzystuje technologię fal radiowych o niskiej mocy, jednakże większej niż Bluetooth. Typowa sieć WiFi liczy kilkanaście urządzeń. Każdy z użytkowników wysyła paczkę danych, według wcześniej zadeklarowanego protokołu (najczęściej jest to protokół TCP/IP), a nośnikiem przesyłanych informacji są fale radiowe. Najczęściej sieć WiFi służy do komunikacji kilku urządzeń będących w jednej sieci oraz podłączeniu ich do Internetu (poprzez punkt dostępu, np. router).

Bezprzewodowa sieć WiFi dzieli się na kilka typów:

1. WiFi - 802.11b. Jest to jedna z serii technologii bezprzewodowych, która obsługuje przepustowość do 11Mb/s. Sygnał tej serii wykorzystuje spektrum częstotliwości

około 2,4GHz. Zasięg sygnału jest stosunkowo duży i z reguły wystarczający w przypadku urządzeń gospodarstwa domowego.

2. WiFi - 802.11g. Jest to standard, który obsługuje transfer danych na poziomie 54Mb/s i wykorzystuje również częstotliwość 2,4GHz, dlatego zasięg jest porównywalny ze standardem 802.11b.
3. WiFi - 802.11n. Jest to najbardziej popularny standard technologii WiFi. Kluczową zmianą jest zastosowanie większej liczby pasm oraz anten (zamiast jednej). Przełożyło się to na przepustowość ponad 100Mb/s i zwiększoną intensywność sygnału.

Większość urządzeń oraz dostępnych modułów komunikacji WiFi, z ludzkie obcuja każdego dnia posiada wsparcie dla wszystkich powyżej opisanych standardów (Agarwal, 2018).

W projekcie urządzenia komunikacji głosowej zasadnym byłoby zastosowanie technologii WiFi Direct, która umożliwia bezpośrednie połączenie urządzeń WiFi, dzięki czemu drukowanie, udostępnianie, synchronizacja i wyświetlanie danych na różnych urządzeniach jest proste i wygodne. Połączenie urządzeń z certyfikatem WiFi Direct wymaga odpowiedniego uwierzytelnienia. Połączenia chronione są za pomocą WPA2 (sposób szyfrowania danych oparty na kluczu, hasła od 8 do maksymalnie 63 znaków). WiFi Direct nie wymaga punktu dostępu, po prostu ustalane jest bezpośrednie połączenie między urządzeniami. Urządzenia z certyfikatem Wi-Fi Direct mogą być używane do różnych rodzajów aplikacji takich jak udostępnianie dowolnych treści, synchronizacja danych, odtwarzanie audio i wideo (Thorn, 2014).

Najlepszym sposobem porównania technologii komunikacji bezprzewodowej jest porównanie parametrów najpopularniejszych modułów komunikacyjnych używanych w aplikacjach urządzeń ubieralnych. Kluczowe parametry zostały przedstawione w tabeli 3.1. Moduł HC-05 reprezentuje klasyczny standard Bluetooth w wersji 2.0. Posiada stosunkowo dużą przepustowość danych (do 3Mb/s), przy niskim, ale stałym poborze prądu na poziomie 40mA. Moduł HM-10 z Bluetooth Low Energy w wersji 4.0 przy niższej prędkości transmisji zużywa prąd do 20mA (tylko w momencie parowania i transmisji). Wadą HC-05 jest brak możliwości podłączenia kilku urządzeń tego typu do jednostki nadrzędnej (np. telefonu komórkowego). Moduł ESP8266 reprezentuje znacznie wyższy poziom przepustowości danych, jednak przy znaczącym poborze energii, osiagającym do 300mA. Komunikacja WiFi dla modułu ESP8266 opiera się o standardowy protokół TCP/IP, a także posiada możliwość zabezpieczenia WPA2.

Tab. 3.1. Zestawienie parametrów popularnych modułów komunikacji bezprzewodowej (opracowanie własne na podstawie (Botland, 2019a, 2019b, 2019c))

	<b>Bluetooth v2.0</b>	<b>Bluetooth v4.0 (LE)</b>	<b>WiFi 802.11n/g/n</b>
	<b>HC-05</b>	<b>HM-10</b>	<b>ESP8266</b>
Przepustowość	3Mb/s	1Mb/s	100Mb/s
Zużycie energii	do 40mA	przy pracy do 20mA	ciągłe 300mA
Zasięg	ok. 10m	ok. 10m	ok. 100m
Bezpieczeństwo danych	parowanie z urządzeniem	parowanie z urządzeniem	WPA2
Kontrola transmisji	brak	brak	TCP/IP
Możliwość podłączenia kilku urządzeń	nie	tak	tak

## 3.4. Klasyfikacja gestów

### 3.4.1. Wprowadzenie

Człowieka niepełnosprawnego można określić jako osobę, która w obliczu życiowej dysfunkcji, do której zaliczyć można uraz, chorobę czy też wadę wrodzoną, nie potrafi wykonywać w zbliżonym stopniu zaawansowania takich samych czynności umysłowych lub fizycznych co osoba, która znajduje się w podobnym wieku (Sękowska, 1985, str. 14). W przypadku osób niepełnosprawnych, dla których komunikowanie się stanowi barierę ze względu na fakt, iż proces ten jest utrudniony lub niemożliwy do osiągnięcia w sposób standardowy, stosowane są metody komunikacji wspomagającej i alternatywnej (AAC – ang. *Alternative and Augmentative Communication*) (Bielak, 2014, str. 26).

### 3.4.2. Metody komunikacji wspomagającej i alternatywnej

Główną cechą charakterystyczną komunikacji alternatywnej jest stosowanie w niej niewerbalnych metod porozumiewania się, co stanowi swoisty zamiennik kodu fonicznego, a więc mowy. Gdy osoba niepełnosprawna do komunikowania się wykorzystuje różnego rodzaju pomoce, jak np. syntezytor mowy czy komputer, to można to określić mianem komunikacji alternatywnej wspomaganej. Jeśli natomiast taka osoba jest samodzielnym i jedynym nośnikiem informacji, wówczas można mówić o komunikacji alternatywnej niewspomaganej. Spośród metod AAC wymienić można następujące grupy systemów znaków:

- system znaków graficznych,

- system znaków przestrzenno-dotykowych,
- system znaków manualnych (Bielak, 2014, str. 26-27).

W ostatniej z wymienionych metod AAC informacja przekazywana jest przy pomocy ruchów rąk oraz palców. W systemie znaków manualnych wyróżnić można kilka podstawowych form przekazu informacji:

1. Alfabet palcowy. Jest to inaczej daktylografia, pozwala na przekazywanie jedynie prostych symboli w postaci liter, liczb i znaków dodatkowych wykorzystując do tego celu ruch i ułożenie palców.
2. Metoda Cued Speech. Polega ona na odczytywaniu wypowiedzi z ust przy jednoczesnym obserwowaniu ruchu jednej ręki. Obserwując poruszające się usta, dostrzega się jednocześnie zsynchronizowane ruchy kończyny oraz układ palców, co pozwala na lepsze rozpoznanie kolejnych głosek.
3. System językowo-migowy (SJM). Wykorzystywany jest tutaj sztuczny język, gdzie treść komunikatu stanowi naturalny szkielet logiczno-gramatyczny języka narodowego, natomiast przekazywany jest przy pomocy gestów języka migowego oraz alfabetu palcowego.
4. Naturalny język migowy. Jest to język mniejszości kulturowej i językowej populacji głuchych. W Polsce wykorzystywany jest Polski Język Migowy (PJM) (Bielak, 2014, str. 29-30).

Naturalny język migowy, w przeciwieństwie do systemu językowo-migowego, posiada własną gramatykę pozycyjną, w której sens zdania określany jest na podstawie kolejności odwzorowywanych wyrazów-gestów. Powstanie takiej formy gramatyki uwarunkowane zostało faktem, iż znaki ideograficzne posiadają stałą formę bez wszelkich odmian. Z tego względu struktury gramatyczne musiały zostać wypracowane w innej formie. Przy odmiennych formach gramatycznych języków fonicznych i migowych, w przypadku zaistnienia potrzeby przetłumaczenia komunikatu z języka narodowego na język migowy lub na odwrót, nie można tłumaczyć słowa po słowie, lecz całe zdania zgodnie z regułami gramatycznymi języka, na który się tłumaczy (Mieszkowicz, 2012).

W przypadku naturalnego języka migowego, wykorzystującego gramatykę pozycyjną, rzeczownik powinien zostać pokazany przez migającego jako pierwszy, a jego klasyfikator bezpośrednio po nim. Może on przyjmować formę przymiotnika, przysłówka, czasownika, zaimka lub też przyimka. W języku migowym wyróżnić można następujące typy klasyfikatorów:

- klasyfikator semantyczny – pełni funkcję zaimka, który zastępuje rzeczownik,

- klasyfikator opisujący – wykorzystywany do opisu kształtów, wielkości czy wzorów,
- klasyfikator instrumentalny – ułożenie dłoni opisuje sposób w jaki używa się dany obiekt,
- klasyfikator elementu – wykorzystuje zarówno ułożenie dłoni jak i ich ruchy, w celu zobrazowania właściwości i ruchów takich elementów jak ogień, woda czy powietrze,
- klasyfikator lokacyjny – może występować w dwóch formach:
  - lokacji – wskazuje lokalizację danego elementu lub też pozycję, względem której ten element się znajduje,
  - ścieżki – pokazuje ścieżkę poruszania się obiektu, jego ruch lub też dystans,
- klasyfikator ciała – wykorzystywany do wskazania na konkretną część ciała (najczęściej w górnej partii),
- klasyfikator części ciała – podobnie jak klasyfikator ciała, również wskazuje konkretną część ciała, z tymże poza obszarem migania,
- klasyfikator pluralny – stanowi liczbę mnogą rzeczownika lub opisywanego podmiotu (Handspeak, 2019).

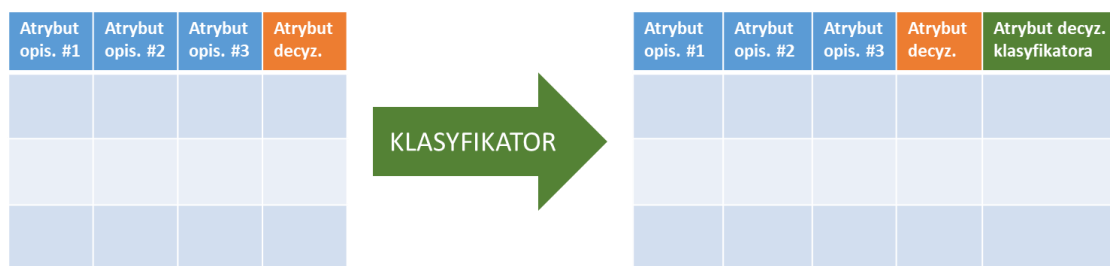
### 3.4.3. Klasyfikatory

Klasyfikacja polega na przypisaniu każdemu zbiorowi danych konkretnej etykiety klasy. W procesie funkcji klasyfikacyjnej (klasyfikatora) można wyróżnić dwa podstawowe etapy:

- budowa klasyfikatora, który opisywać będzie predefiniowany zbiór klas,
- wykorzystanie stworzonego klasyfikatora w procesie odwzorowywania danych wejściowych w zbiór klas (klasyfikacja) w dwóch fazach: testowania oraz predykcji.

Dane wejściowe to zbiór krotek (przykładów), które są wartościami atrybutów opisowych i odpowiadających im atrybutom decyzyjnym. Klasyfikator odpowiedzialny jest za predykcję atrybutu decyzyjnego dla danego zbioru danych, dla których wcześniej wartość atrybutu decyzyjnego nie była znana (Diks, 2007).

Proces tworzenia klasyfikatora wymaga przygotowania dwóch zbiorów danych: zbioru uczącego oraz zbioru testującego, tak jak przedstawiono na rysunku 3.2. Zbiór uczący stanowi podstawę dla stworzenia klasyfikatora, natomiast zbiór testujący weryfikuje dokładność klasyfikatora w przyporządkowywaniu danych do właściwych klas.



Rys. 3.2. Proces tworzenia klasyfikatora (opracowanie własne)

W momencie, gdy klasyfikator został odpowiednio zamodelowany i sprawdzony na zbiorze testującym, można ocenić jego dokładność. Znane wartości atrybutów decyzyjnych dla zbioru testującego porównywane są z wartościami uzyskanymi przez zastosowanie klasyfikatora. Na tej podstawie można określić współczynnik dokładności, według wzoru 3.1.

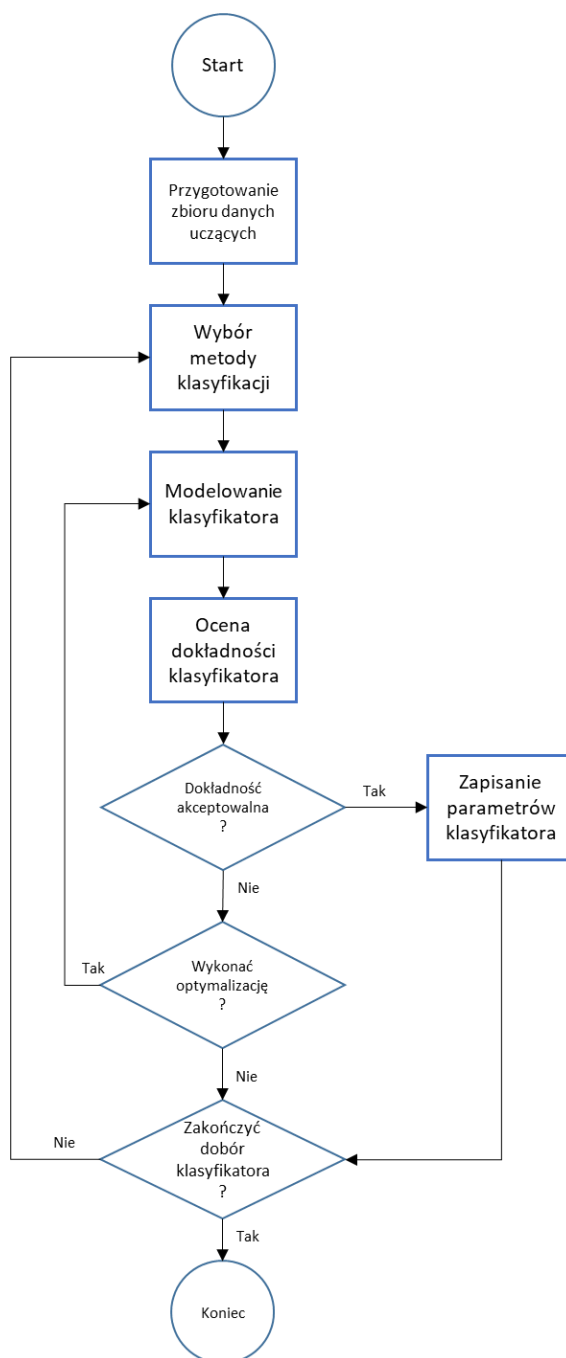
$$\text{dokładność} = \frac{\text{liczba poprawnie zaklasyfikowanych przykładów}}{\text{liczba wszystkich przykładów}} \cdot 100\% \quad (3.1)$$

Jeśli zamodelowany klasyfikator nie spełnia założonych wymagań dokładności, należy przeprowadzić jego optymalizację lub podjęcie próby zastosowania innej metody klasyfikacji dla danego zbioru atrybutów opisowych. Proces ten przedstawiony został na rysunku 3.3. Jeśli osiągnięty poziom dokładności klasyfikatora jest akceptowalny, wówczas można rozpocząć proces klasyfikacji nowego zbioru danych, dla którego zbiór atrybutów decyzyjnych nie jest znany, w celu przyporządkowania do odpowiednich klas (Diks, 2007).

Ze względu na to, iż istnieje wiele metod klasyfikacji, stworzone zostały kryteria porównawcze, dzięki którym można lepiej ocenić przydatność danego klasyfikatora i porównać go z pozostałymi. Można tutaj wyróżnić:

- dokładność predykcji – zdolność modelu do prawidłowego przewidzenia wartości klasy nowego przykładu,
- efektywność – koszt obliczeniowy, który jest wygenerowany w procesie tworzenia i stosowania klasyfikatora,
- odporność modelu – zdolność modelu do prawidłowego przewidzenia wartości klasy w momencie posiadania niepełnych danych lub występujących szumów w atrybutach opisujących,
- skalowalność – możliwość stworzenia modelu klasyfikatora na nowych i dowolnie dużych zbiorach danych,

- interpretowalność – stopień zrozumienia mechanizmów determinujących działanie klasyfikatora (Diks, 2007).



Rys. 3.3. Schemat procesu klasyfikacji (opracowanie własne)

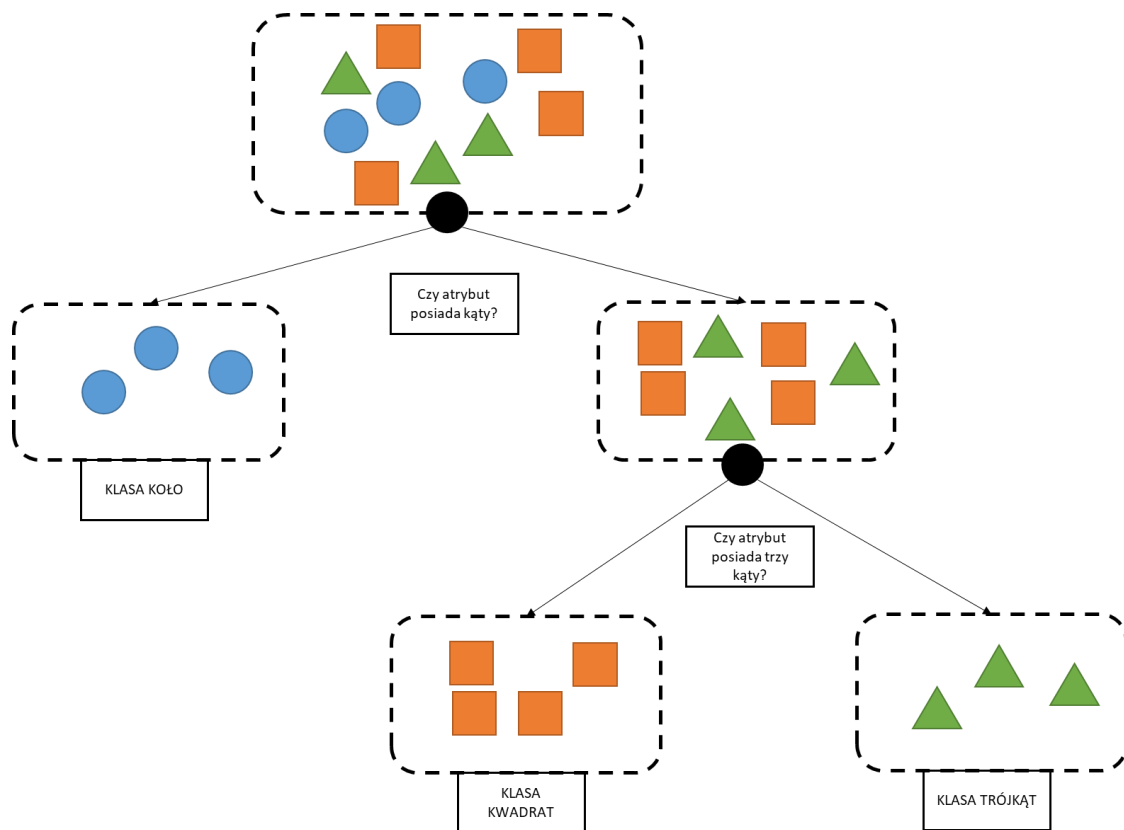
### 3.4.4. Metody klasyfikacji

Istnieje wiele metod klasyfikacji, z których każda posiada inny schemat działania, mogący stanowić o wadach i zaletach takiej metody w odniesieniu do aktualnie stawianego problemu



klasyfikacyjnego. Poniżej zaprezentowano skrótowo ideę kilku sposobów modelowania klasyfikatorów, które zostały przetestowane w niniejszej pracy:

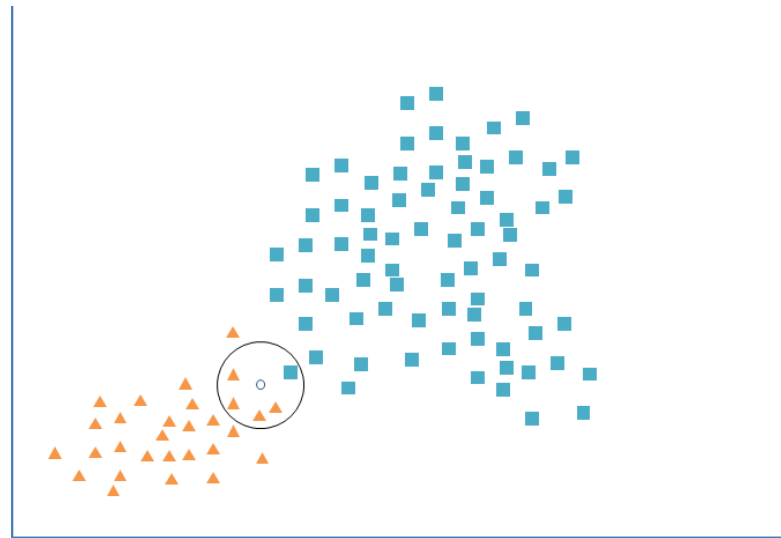
1. Drzewa decyzyjne. Są one skierowanym, acyklicznym grafem, posiadającym strukturę drzewiastą, gdzie każdy z wierzchołków stanowi elementarny test wykonywany na atrybucie, każdy łuk jest wynikiem tego testu, natomiast każdy liść przedstawia pojedynczą klasę lub przedział wartości klas (rysunek 3.4). Podział dokonywany jest do momentu, aż w zbiorze danych nie pozostaną atrybuty należące tylko do jednej klasy (Diks, 2007).



Rys. 3.4. Obrazowy przykład działania drzewa decyzyjnego (opracowanie własne)

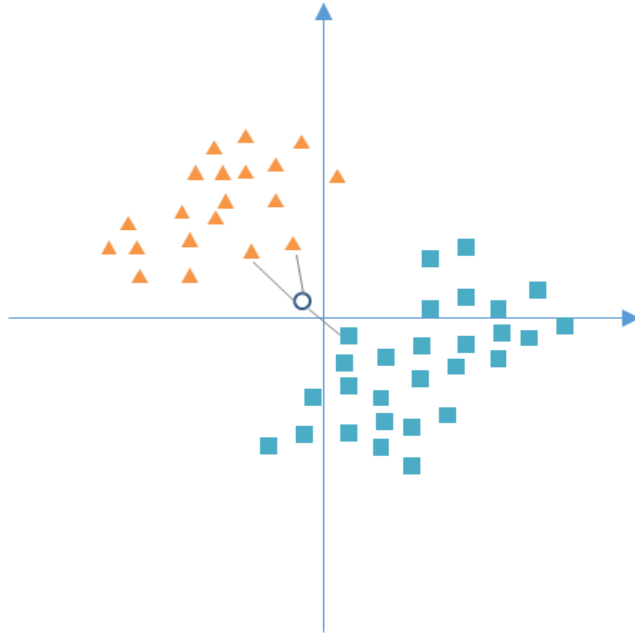
2. Klasyfikator Bayesa. Określa (szacuje) z jakim prawdopodobieństwem dany obiekt, o określonych atrybutach opisowych, należy do konkretnej klasy decyzyjnej i przypisuje do tego obiektu najbardziej prawdopodobną klasę (Grzyb, 2007). Na rysunku 3.5 widać dwa typy obiektów, należących do dwóch klas: kwadrat oraz trójkąt. Ze względu na to, iż obiektów typu kwadrat jest trzy razy więcej niż obiektów typu trójkąt, w analizie Bayesowskiej można założyć, że nowy obiekt posiadać będzie trzykrotnie wyższe prawdopodobieństwo na przynależenie do klasy kwadrat niż do klasy trójkąt. Takie prawdopodobieństwo nazywane jest *a priori*.

Prawdopodobieństwem *a posteriori* w tym przypadku może być prawdopodobieństwo przynależności nowego obiektu do tej klasy, której większa liczba reprezentantów znajduje się w zadanej odległości od obiektu. Na rysunku 3.5 widać, że, w zadany okręgu, obiektów klasy trójkąt jest czterokrotnie więcej niż obiektów klasy kwadrat. Ostatecznie nowy przykład zostanie przyporządkowany do klasy trójkąt, ze względu na to iż większe jest tutaj prawdopodobieństwo *a posteriori* do takiej przynależności (Harasymczuk, 2019).



Rys. 3.5. Przykład działania klasyfikatora Bayesa (opracowanie własne na podstawie (Harasymczuk, 2019))

3. Klasyfikator najbliższego sąsiedztwa. Jego działanie polega na przyporządkowywaniu obiektowi takiej klasy, dla której jej liczba reprezentantów jest największa w zadanej odległości (rysunek 3.6). Problemem wynikającym ze stosowania tej metody jest możliwość posiadania różnej skali, jednostek lub też przedziałów zmienności przez kolejne atrybuty opisowe. W takim przypadku, gdyby zdecydowano się zastosować funkcję odległości bez wcześniejszych przekształceń, mogłoby dojść do sytuacji, w której jeden z atrybutów stałby się dominujący, redukując wpływ na wynik pozostałych atrybutów. Aby nie dopuścić do takiej sytuacji, należy przeprowadzić normalizację atrybutów i przyporządkować im odpowiednie wagi (Andrzejewski, 2008).



Rys. 3.6. Przykład działania klasyfikatora najbliższego sąsiedztwa (opracowanie własne na podstawie (Harasymczuk, 2019))

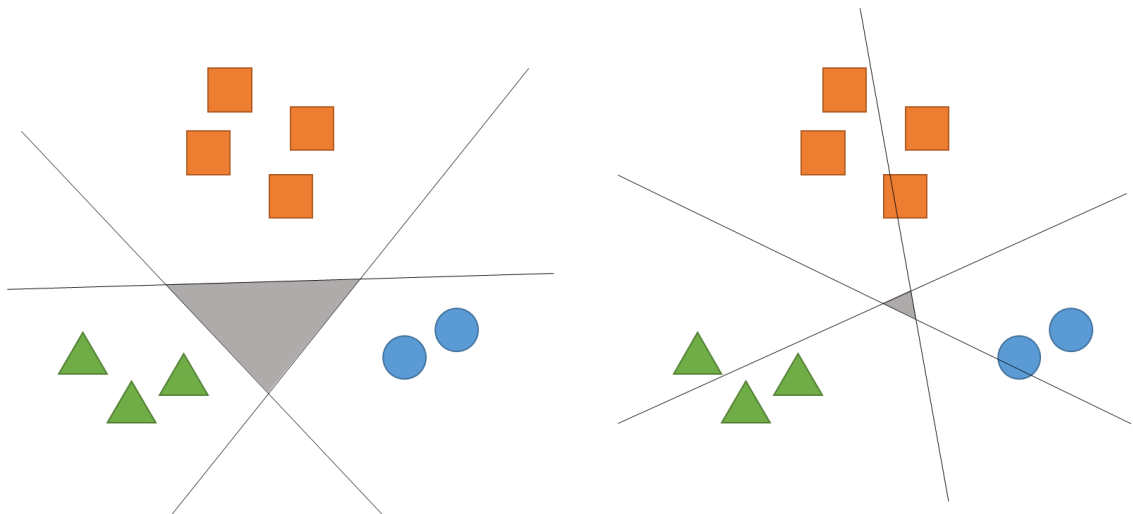
### 3.4.5. Strategie rozwiązywania problemów wieloklasowych

W przypadku klasyfikacji problemu binarnego, klasyfikator musi przyporządkować dany przykład do jednej z dwóch klas, jednak złożoność problemu klasyfikacji rośnie wraz ze wzrostem liczby docelowych klas. Dwoma najpopularniejszymi strategiami rozbicia problemu klasyfikacji wieloklasowej na przypadki binarne są:

1. Strategia jeden przeciw pozostałym – OVA (ang. *One Versus Rest*). Strategia ta polega na utworzeniu dla  $N$  klas  $N-1$  klasyfikatorów i przetestowaniu każdego przykładu przy pomocy wszystkich utworzonych klasyfikatorów. W sytuacji idealnej, tylko jeden z klasyfikatorów przyporządkuje przykład do klasy, jednak w praktyce może okazać się, że więcej próbek znajdzie swoje odwzorowanie w zbiorze klas. W drugim z opisanych przypadków można przyjąć strategię odrzucenia danego przykładu i nieprzyporządkowanie go do żadnej z klas, co polepszyłoby dokładność klasyfikacji, jednak spowodowałoby zwiększenie współczynnika odrzucenia. Innym podejściem byłoby określenie stopnia (dokładności) przyporządkowania do danej klasy i wybranie klasy z najlepszym dopasowaniem.
2. Strategia jeden przeciw jednemu – OVO (ang. *One Versus One*). W tej strategii tworzone jest  $N * (N-1)/2$  klasyfikatorów binarnych, porównujących każde dwie pary klas z kolejną próbką. Podobnie jak wcześniej, trudnym do spełnienia, idealnym przypadkiem byłoby przypisanie do próbki docelowej klasy  $N-1$  razy, co zwiększyłoby dokładność i zminimalizowałoby błędne przypisania, jednak zwiększyłoby

współczynnik odrzucenia. W praktyce jednak docelowa klasa wybierana jest na podstawie największej liczby dopasowań (Chmielnicki, 2012).

Na rysunku 3.7 przedstawiono porównanie strategii OVR i OVO. Zacięnięwany trójkątny obszar oznacza tzw. „obszar niepewności”, dla którego próbka może nie zostać przypisana do żadnej z klas, jeśli znajdzie się w tym obszarze (w strategii OVR) lub może zostać błędnie sklasyfikowana (w strategii OVO). W przypadku strategii jeden przeciw jednemu, należy zamodelować  $N * (N-1)/2$  klasyfikatorów, co znacząco podnosi koszt obliczeniowy w porównaniu z pierwszą strategią. Przy zastosowaniu strategii OVO istnieje lepsze zbalansowanie klas przeciwstawnych, a posiadając klasy reprezentowane przez zbliżoną liczbę próbek, w procesie uczenia klasyfikatora obecne będzie równomierne zbilansowanie reprezentacji klas. W przypadku strategii OVR na etapie modelowania klasyfikatora może łatwo dojść do efektu nadreprezentacji jednej z klas (Chmielnicki, 2012).



Rys. 3.7. Porównanie strategii OVR i OVO (opracowanie własne na podstawie (Chmielnicki, 2012))

## 3.5. Analiza istniejących rozwiązań

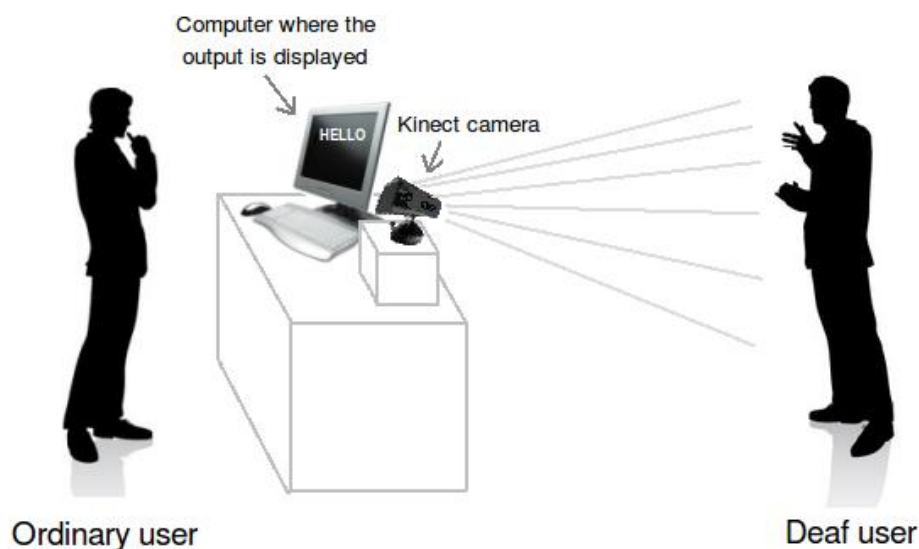
### 3.5.1. Wprowadzenie

Rozwijająca się technologia otwiera wiele możliwości dla tworzenia nowych urządzeń, dlatego problem rozpoznawania gestów został podjęty przez wielu naukowców. Dostępne urządzenia, które potrafią interpretować różne gesty można podzielić na dwie kategorie. Jedną z nich są systemy wizyjne oparte na analizie obrazu z kamery, natomiast drugą są urządzenia elektroniki ubieralnej (ang. *wearables*), które posiadają zabudowany system mikroprocesorowy z układami odpowiednich sensorów.

### 3.5.2. Systemy wizyjne

Jednym z najpopularniejszych systemów wizyjnych rejestrujących ruch jest Kinect – urządzenie wejściowe dla konsoli Xbox wyprodukowane przez firmę Microsoft. Działanie Kinecta opiera się na dwóch kamerach, promienniku podczerwieni, akcelerometrze i napędzie pozwalającym na uchylanie głowicy. Jedna z kamer rejestruje naturalny obraz, który jest przetwarzany i wykorzystywany do rozpoznania twarzy gracza. Kluczowym elementem, który umożliwia sterowanie za pomocą gestów jest kamera podczerwieni. Kamera emituje światło podczerwone w wielu wiązkach, które, po odbiciu od elementów znajdujących się w pomieszczeniu, wracają. Czas wędrówki każdego z emitowanych promieni podczerwonych jest rejestrowany i na tej podstawie tworzony jest obraz zachowujący głębię przedmiotów oraz osób. Zaawansowane oprogramowanie urządzenia Kinect, pozwala rozpoznawać graczy i rejestrować gesty całego ciała. Urządzenie śledzące rejestruje w czasie rzeczywistym osobno 48 części rozpoznanej sylwetki człowieka. Kinect oprócz gier znalazł swoje zastosowania w środowiskach zajmujących się przetwarzaniem obrazów. Stanowi zamiennik dla drogich sensorów laserowych. Istnieje wiele aplikacji i bibliotek pozwalających na obsługę sensora Kinect przez wydajne komputery na wielu systemach operacyjnych (Bednarek, 2014).

W roku 2013 grupa naukowców postanowiła wykorzystać potencjał Kinecta. Stworzono prototyp urządzenia, które tłumaczy język migowy na język narodowy. KinectTranslator, bo tak nazywa się to rozwiązanie, odczytuje ruchy osoby migającej i na podstawie wcześniej zdefiniowanych słów, znajdujących się w bazie danych, tłumaczy rozmowę w czasie rzeczywistym. Założenia projektu zostały pokazane na rysunku 3.8. Autorzy przy tworzeniu oprogramowania skorzystali z metody sieci neuronowych, wspomagającej uczenie się nowych słów, dzięki czemu bazę słów mogliby tworzyć użytkownicy. Nauczenie aplikacji nowego słowa polegałoby na kilkukrotnym wykonaniu prawidłowego gestu, a następnie wykonaniu kilku nieprawidłowych ruchów, co skutkowałoby kompletnym dostrojeniem parametrów danego słowa lub wyrażenia. Autorzy rozwiązania, w tym osoby głuche posługujące się językiem migowym, planowali stworzyć tłumaczenie z języka narodowego na język migowy oparte na animacji, która potrafiłaby migać zadane słowa. W dostępnych źródłach nie znaleziono informacji czy urządzenie zostało wprowadzone na rynek (Michalak, 2013).

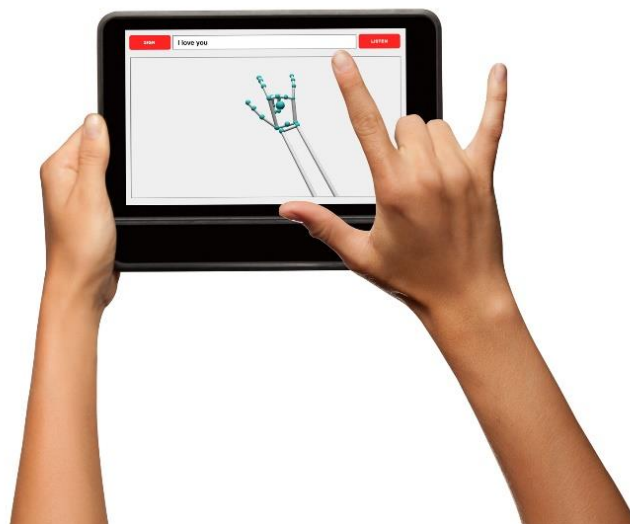


Rys. 3.8. Wizualizacja koncepcji KinectTranslator (Mériaudeau, 2012)

W roku 2014 naukowcy z ETH Zurich dzięki opracowaniu nowego algorytmu, stworzyli aplikację analizującą wykonywane przez użytkownika gesty. Użytkownik wykonuje gesty jedną ręką, w odpowiedniej odległości od kamery telefonu komórkowego, a aplikacja działająca w tle przetwarza obraz i analizuje ruchy użytkownika. Baza odczytywanych gestów jest ściśle ograniczona i nie pozwala na dodanie własnych gestów. Istniejące w bazie gesty pozwalają na sterowanie telefonem komórkowym, tj. przełączanie między kartami przeglądarki, powiększanie lub pomniejszanie skali mapy, przesuwanie stron w książkach. Większość systemów wizyjnych opiera się na komputerach wyposażonych w wydajne procesory i karty graficzne, dlatego rozwiązanie naukowców z Zurichu jest o tyle innowacyjne, że za pomocą procesora telefonu komórkowego jest w stanie, w czasie rzeczywistym, przetwarzać obraz (Sciencedaily, 2014).

W 2014 roku 9-osobowa grupa z Rochester Institute of Technology w Nowym York'u stworzyła urządzenie, wspomagające proces komunikacji pomiędzy osobami niesłyszącymi i niedosłyszącymi a słyszącymi. Produkt MotionSavvy's, nazywany UNI, to tablet z funkcją rozpoznawania ruchów rąk (rysunek 3.9), wykorzystujący moduł Leap Motion (kilka kamer oraz czujnik światła IR, potrafiący wykrywać ruch w przestrzeni 3D, śledzący ruchy obu rąk oraz 10 palców z dużą prędkością i precyzją (Botland, 2019d)). UNI umożliwił płynną, dwukierunkową komunikację między głuchym a słyszącym za pomocą dwóch różnych technologii. Urządzenie MotionSavvy's, monitorowało gesty wykonywane w języku migowym i przekształcało je na tekst wyświetlany na ekranie tabletu. Następnie, gdy osoba słysząca odpowiadała, aplikacja analizowała głos i zamieniała go na tekst, który również wyświetlany był w oknie historii konwersacji. Jednym z wyzwań, które postawiło sobie MotionSavvy było stworzenie cyfrowego słownika dla większości najpopularniejszych języków migowych.

Problem był trudny do rozwiązania, ponieważ języki migowe na świecie różnią się nie tylko gestami, ale także gramatyką. Zespół MotionSavvy szukał kompleksowego sposobu na dodawanie gestów do cyfrowego słownika przez samych użytkowników. W 2014 roku firma miała 300 gestów w bazie danych i zapowiadała, że w kolejnym roku w słowniku znajdzie się do 50 000 gestów. Urządzenie miało kosztować 198\$ (ok. 770zł). Aktualnie w dostępnych źródłach internetowych nie znaleziono informacji czy urządzenie zostało wprowadzone do sprzedaży (Strauss, 2014).



Rys. 3.9. MotionSavvy's UNI, urządzenie do rozpoznawania gestów w języku migowym (Medgadget, 2014)

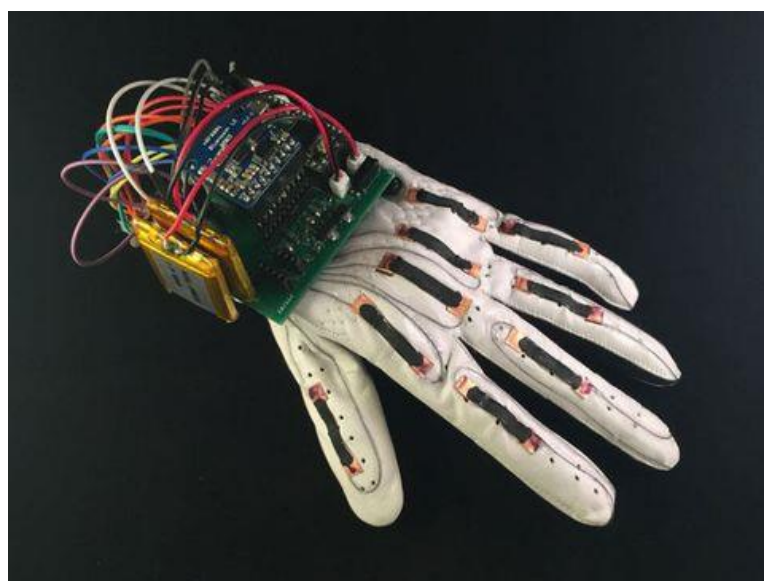
### 3.5.3. Urządzenia wearable

W roku 2016 dwóch studentów z Uniwersytetu w Waszyngtonie pracowało nad stworzeniem nowego sposobu komunikacji między osobami niesłyszącymi a słyszącymi. Stworzyli rękawice SignAloud (rysunek 3.10), które tłumaczyły amerykański język migowy (ASL) na mowę i tekst. Urządzenie opierało się na dwóch bezprzewodowych rękawicach komunikujących się za pomocą Bluetooth z komputerem, który przetwarzał i analizował dane. Rękawice wyposażyli w czujniki ugięcia palców, akcelerometr, mikroprocesor, baterię oraz moduł Bluetooth. Komputer sprawdzał dane pod kątem gestów poprzez różne sekwencyjne regresje statystyczne, podobne do sieci neuronowej. Gdy dany gest został poprawnie pokazany, określone słowo lub wyrażenie było odtwarzane przez głośnik. Autorzy zapewniali, że ich rękawice są lekkie, bardzo wygodne i nie powodują dyskomfortu przy użytkowaniu. W dostępnych źródłach nie znaleziono finalnej wersji produktu (Wanshel, 2016).



Rys. 3.10. SignAloud, prototyp rękawicy studentów z Uniwersytetu w Waszyngtonie (Perna, 2017)

W roku 2017 roku grupa naukowców z Uniwersytetu Kalifornijskiego w San Diego zbudowała rękawicę elektroniczną (rysunek 3.11), która wykrywała znaki używane w amerykańskim języku migowym (ASL) i tłumaczyła na litery w alfabecie angielskim. Rękawica wykorzystywała czujniki ugięcia palców do śledzenia kąta ich zgięcia oraz akcelerometr do śledzenia pozycji dłoni w przestrzeni. Kompilacja tych sensorów pozwoliła naukowcom dokładnie określić orientację dłoni i pozycję palców. Stworzona przez nich rękawica potrafiła rozróżnić wszystkie 26 liter alfabetu ASL. Naukowcy zbudowali prototyp za mniej niż 100\$ (390zł). Autorzy planowali w przyszłości dodać możliwość rozpoznawania pełnych wyrażen, jednak w dostępnych źródłach nie znaleziono informacji o kolejnych postępach w funkcjonalności ich produktu (Thompson, 2017).



Rys. 3.11. Prototyp rękawicy naukowców z San Diego (Thompson, 2017)



### 3.5.4. Porównanie rozwiązań

Po dzień dzisiejszy stworzonych zostało kilka zaawansowanych technologicznie urządzeń, których głównym zdaniem jest przetwarzanie gestów języka na tekst. Każda z grup naukowców starała się podejść do tematu w inny sposób, stosując różne technologie i metody klasyfikacji gestów, dlatego stworzone prototypy posiadają różne wady i zalety, które zostały przedstawione w tabeli 3.2 Stworzenie zaawansowanych układów elektronicznych, wspierających tłumaczenie języków migowych, spotkało się z pozytywnym odbiorem ze strony wielu osób, jednak opisane urządzenia nie zostały wprowadzone na rynek.

Tab. 3.2. Porównanie popularnych urządzeń rozpoznających gesty (opracowanie własne)

	WIZYJNE		WEARABLES	
	Kinect Translate	MotionSavvy's UNI	SignAloud	Rękawica UC w San Diego
Konieczność użycia PC	tak	nie	tak	nie
Urządzenie mobilne	nie	tak	tak	tak
Komunikacja bezprzewodowa	nie	tak	tak	tak
Rozpoznawanie liter	tak	tak	tak	tak
Rozpoznawanie słów	tak	tak	tak	nie
Wygoda stosowania	tak	tak	tak	nie
Możliwość uczenia gestów	tak	nie	nie	nie
Koszt urządzenia	wysoki	wysoki	średni	niski

KinectTranslate okazało się bardzo efektywnym rozwiązaniem, analiza obrazu była bardzo dokładna, a zastosowanie wydajnego komputera pozwalało bardzo szybko analizować dane z sensora. Zastosowanie Kinect było o tyle komfortowe w użytkowaniu, że nie wymagało noszenia specjalnych rękawic lub kontrolerów, jednak ta technologia niosła za sobą ograniczenia. Urządzenia można było używać wyłącznie stacjonarnie. Potrzebna była odpowiednio duża przestrzeń, aby sensor mógł widzieć daną osobę. Poza tym zastosowanie wydajnego komputera wiązało się z dużymi kosztami dla potencjalnego użytkownika (Michalak, 2013).

MotionSavvy's swoje działanie opierało na tablecie, a więc urządzeniu mobilnym, co było niewątpliwą zaletą. Zastosowanie modułu Leap Motion pozwalało bardzo efektywnie rejestrować orientację dłoni w przestrzeni. Dodając do tego bogaty słownik urządzenie

mogłoby być bezkonkurencyjne na rynku. Jednak główną wadą urządzenia był bardzo wysoki koszt całościowy, związany z zastosowaniem odpowiedniego modelu tabletu oraz wspomnianego wyżej modułu czujników (Strauss, 2014).

SignAloud oraz rękawica stworzona przez naukowców z San Diego, były podobnymi urządzeniami – bazowały na bardzo zbliżonych podzespołach. Różniły się tylko zastosowaną jednostką centralną. SignAloud analizował dane za pomocą komputera, rękawica z uniwersytetu kalifornijskiego wysyłała dane do telefonu komórkowego, który rozróżniał gesty palców. Pomimo tego, że obie jednostki były mobilne, korzystanie z telefonu komórkowego jest wygodniejsze. Wiąże się ono jednak z ograniczeniem wydajności obliczeń, co powoduje, że rękawica naukowców z San Diego rozpoznawała tylko litery alfabetu. Oba rozwiązania były jednak rozwiązaniami nisko budżetowymi (Wanshel, 2016; Thompson, 2017).

Reasumując, można stwierdzić, że przy obecnej sytuacji na rynku, nie istnieje urządzenie, które potrafiłoby kompleksowo tłumaczyć język migowy na tekst czy mowę, zachowując przy tym mobilność i pełną swobodę ruchów użytkownika. Zasadne jest zatem podjęcie próby stworzenia prototypu urządzenia, wypełniającego tę lukę, ponieważ mogłoby ono zrewolucjonizować sytuację osób głuchoniemych na rynku pracy i przyczynić się do ich aktywizacji zawodowej m.in. w środowiskach produkcyjnych.

## 4. Metodyka pracy

### 4.1. Koncepcja i plan prac

Proces tworzenia prototypu urządzenia w dziedzinie, która została w stosunkowo nieznacznym stopniu przebadana stanowi wyzwanie o wysokim stopniu trudności. Złożoność obliczeniowa oraz zależności jakie występują w porozumiewaniu się przez osoby niesłyszące i słabosłyszące determinują środek ciężkości projektu, który obejmuje wiele aspektów. Aby skutecznie przeprowadzić cały proces, począwszy od wypracowania koncepcji a skończywszy na poprawnym rozpoznawaniu gestów, wymagane było przygotowanie szczegółowego planu. Jego elementy zostały zaprezentowane w tabeli 4.1.

Podstawą do wykonania w pełni funkcjonalnego urządzenia do komunikacji głosowej dla osób głuchoniemych w warunkach produkcyjnych było przeprowadzenie obszernej analizy literaturowej, w celu zdobycia niezbędnej wiedzy z zakresu posługiwania się językiem migowym, metod klasyfikacji czy tworzenia aplikacji mobilnych. Poznanie istniejących rozwiązań w obszarze komunikacji osób głuchoniemych z osobami słyszącymi i metod interpretacji języka migowego stało się istotnym czynnikiem do określenia kierunku w jakim projekt powinien się rozwinąć. Kluczowym punktem tego etapu było ustalenie wstępnej koncepcji urządzenia. Pozwoliło to na zgłębienie literatury ze znacznie węższego zakresu i weryfikację danych technologii pod kątem funkcjonalności i możliwości implementacyjnych.

Etap weryfikacji dostępnych technologii pozwolił na ostateczne określenie jakie technologie zostaną wykorzystane w projekcie. Z najbardziej popularnych rozwiązań komunikacji bezprzewodowej, a więc Bluetooth oraz WiFi, zostały zebrane wszystkie aspekty przemawiające za i przeciw danej technologii. Zestaw kryteriów został dostosowany na podstawie wcześniejszej analizy istniejących rozwiązań i trudności z jakimi autorzy innych prac się zmagali, np. szybkość transmisji danych, pobór energii i inne. Pod uwagę wzięte zostały również możliwości konkretnych sensorów i mikrokontrolerów pod kątem przydatności w ekstrakcji wartościowych atrybutów opisowych gestów.

Stworzenie założeń konstrukcyjnych zdefiniowało ostateczny dobór technologii komunikacyjnych, sensorów, mikrokontrolera oraz pozostałych podzespołów. Ponadto rozplanowany został układ fizyczny rękawicy, a więc rozmieszczenie wszystkich komponentów i połączenia elektryczne między nimi.

Tab. 4.1. Plan prac wykonania prototypu rękawicy dla osób głuchoniemych (opracowanie własne)

Działanie	Opis
Analiza literatury	<ul style="list-style-type: none"> <li>• zapoznanie się z istniejącymi rozwiązaniami w dziedzinie komunikacji osób głuchoniemych i wykorzystywanymi technologiami,</li> <li>• wstępne opracowanie koncepcji na podstawie istniejących rozwiązań.</li> </ul>
Weryfikacja dostępnych technologii	<ul style="list-style-type: none"> <li>• porównanie technologii komunikacyjnych: Bluetooth, WiFi,</li> <li>• porównanie metod zbierania i selekcji danych: czujniki nacisku a czujniki ugięcia, dane z akcelerometru a żyroskopu.</li> </ul>
Stworzenie założeń konstrukcyjnych	<ul style="list-style-type: none"> <li>• wybór technologii komunikacyjnych,</li> <li>• wybór sensorów i mikrokontrolera,</li> <li>• rozplanowanie układu rękawicy (rozmieszczenie sensorów, mikrokontrolera, modułu komunikacyjnego, baterii).</li> </ul>
Opisanie założeń projektowych	<ul style="list-style-type: none"> <li>• wykonanie podziału zadań pomiędzy członków grupy projektowej,</li> <li>• ustalenie kryteriów akceptowalności działania modułu rozpoznającego gesty.</li> </ul>
Konstrukcja prototypu rękawicy	<ul style="list-style-type: none"> <li>• wykonanie połączeń elektrycznych modułów i sensorów,</li> <li>• wykonanie prototypu w wersji tekstylno-elektronicznej.</li> </ul>
Stworzenie algorytmu pozyskującego dane z sensorów i algorytmu klasyfikującego gesty	<ul style="list-style-type: none"> <li>• określenie mechanizmów zbierania danych,</li> <li>• ustalenie sekwencji działania programu i aplikacji,</li> <li>• dobór atrybutów opisowych i algorytmów klasyfikacyjnych.</li> </ul>
Oprogramowanie rękawicy i aplikacji	<ul style="list-style-type: none"> <li>• oprogramowanie mikrokontrolera rękawicy,</li> <li>• stworzenie aplikacji mobilnej,</li> <li>• przygotowanie bazy słów środowiska produkcyjnego.</li> </ul>
Testy	<ul style="list-style-type: none"> <li>• weryfikacja stworzonych programów.</li> </ul>
Działania optymalizacyjne	<ul style="list-style-type: none"> <li>• wykonanie działań optymalizacyjnych algorytmów,</li> <li>• wykonanie ponownych testów działania programów.</li> </ul>

Kolejnym ważnym etapem było opisanie założeń projektowych. W tej części projektu rozplanowane zostały wszystkie zadania, jakie musiały zostać wykonane aż do ostatniej fazy projektu. Przypisanie konkretnych punktów dla każdego z członków projektu było istotnym elementem zarówno z punktu widzenia obciążenia jak i od strony kolejności wykonywania działań. Ważnym aspektem projektu był bowiem taki podział zadań, aby możliwie precyzyjnie zająć prace nad aplikacją mobilną oraz algorytmami ekstrakcji atrybutów opisowych, aby umożliwić płynne testowanie obu tych części, z których każda była ściśle zależna od drugiej. W tym etapie projektu ustalone zostały także kryteria akceptowalności działania modułu rozpoznającego gesty, a więc przede wszystkim wymagana dokładność i powtarzalność klasyfikacji wykonywanych ruchów. W przypadku, gdyby założone kryteria nie zostały osiągnięte, wówczas należałoby wykonać działania optymalizacyjne algorytmów, zmienić metodę lub strategię klasyfikacji, dobrać inne atrybuty opisowe lub też ponownie zweryfikować dobrane komponenty rękawicy, a w szczególności sensory i dane jakie można z nich uzyskać.

Etap konstrukcji prototypu rękawicy polegał przede wszystkim na przymocowaniu wszystkich komponentów (szycie i klejenie) do materiału rękawicy oraz wykonaniu i zabezpieczeniu połączeń elektrycznych. Następnie przeprowadzone zostały testy wytrzymałościowe konstrukcji na naturalne i gwałtowne ruchy kończyny, a także skalibrowane zostały odczyty z sensorów i zweryfikowana ich powtarzalność, aby uniknąć niepotrzebnych błędów i zakłóceń w kolejnych fazach projektu.

Kolejne etapy pracy związane były ściśle z pracami programistycznymi. Utworzenie algorytmu pozyskującego dane z sensorów i algorytmu klasyfikującego gesty nastąpiło w wyniku licznych testów. Ustalone musiały zostać między innymi odpowiednia częstotliwość pozyskiwania wartości pomiarowych z czujników ugięcia oraz żyroskopu i akcelerometru, liczba odczytów składająca się na jedną ramkę gestu, optymalna liczba ramek, najbardziej charakterystyczne atrybuty opisowe oraz wiele innych parametrów. Metody i strategie klasyfikacyjne zostały dopasowane do obliczanych wartości charakterystycznych również na drodze testów i działań optymalizacyjnych, w których skład wchodziły m.in. korekcje parametrów klasyfikatora czy zmiany w częstotliwości odczytu danych z sensorów. Przy wypracowywaniu sekwencji działania aplikacji mobilnej i mikrokontrolera, szczególny nacisk został położony na szybkość wykonywanych obliczeń oraz na szybkość i przepustowość komunikacji bezprzewodowej, co było niezbędne do zminimalizowania strat w odbiorze danych z sensorów. Po stworzeniu obu programów, należało przygotować bazę słów, które są charakterystyczne i często wykorzystywane w środowisku produkcyjnym. Utworzone algorytmy klasyfikacyjne przetestowane zostały na przygotowanej bazie słów. Po weryfikacji

mierników skuteczności działania klasyfikacji, wykonane zostały ponowne działania optymalizacyjne i testy algorytmów.

## 4.2. Struktura rękawicy oraz oprogramowania

### 4.2.1. Zastosowane komponenty sprzętowe

Analizując istniejące rozwiązania oraz kierując się zdobytym do tego czasu doświadczeniem, dobrano komponenty dla rękawicy w taki sposób, aby spełniała podstawowe założenia związane z wygodą oraz niezawodnością działania. Wszystkie elementy elektroniki zostały połączone na stałe według schematu elektrycznego (załącznik 1).

Szkieletem opisywanego urządzenia jest rękawica. Z uwagi na fakt tworzenia pierwszego prototypu funkcjonalnego wybrana została niskobudżetowa wersja, tj. bawełniana rękawica przeznaczona do prac, w których wymagana jest sprawność manualna (rysunek 4.1). Zaletą takich rękawic jest wysoka wytrzymałość oraz brak występowania zjawiska elektryzowania (posiada właściwości antystatyczne). Drugi aspekt jest kluczowy, ponieważ zastosowana niskonapięciowa elektronika sterowania jest bardzo czuła, na każdy rodzaj zakłóceń przepięciowych. Rękawica dobrze przepuszcza powietrze i jest rozciągliwa, dlatego dłuższe użytkowanie nie powoduje dyskomfortu i łatwo dopasowuje się do rozmiaru dłoni użytkownika (Mirage, 2018).



Rys. 4.1. Bawełniana rękawica (Mirage, 2018)

Najważniejszym ogniwem urządzenia komunikacji głosowej był układ mikroprocesorowy. Przy dokonywaniu wyboru, wzięte zostały pod uwagę następujące czynniki:

1. Niewielki rozmiar. Ma on istotny wpływ na komfort użytkowania. Ważne, aby jednostka była albo bardzo mała (wymiary do 40mm x 40mm) albo posiadała elastyczną płytkę PCB (elektroniczna płytką drukowaną), która mogłaby dopasowywać się do dłoni.

2. Energooszczędność. Był jednym z ważniejszych argumentów z uwagi na fakt, iż urządzenie miało być zasilane niewielką baterią.
3. Wydajna jednostka procesorowa. Oparta jest na procesorze taktującym zegarem o częstotliwości minimum 16MHz oraz pamięci *flash* ponad 150kB.
4. Zintegrowany akcelerometr z żyroskopem. Pozwala odczytywać zmiany położenia układu, a co za tym idzie całej ręki. Zintegrowane sensory pozwalają wyeliminować dodatkowe połączenia, jakie musiałyby być zrealizowane przy podłączeniu dodatkowego modułu peryferyjnego.
5. Zintegrowany moduł komunikacyjny Bluetooth. Był on niezbędny do wymiany danych pomiędzy rękawicą a aplikacją mobilną. Minimalna oczekiwana wersja modułu Bluetooth to 4.0. Jest to jednostka energooszczędna (LE), która pozwala na podłączenie do urządzenia sterującego kliku użytkowników.
6. Możliwość podpięcia pięciu sensorów analogowych. Jest to wymóg dla czujników realizujących odczytywanie poziomu zgięcia palców.
7. Otwarte środowisko programistyczne. Pozwala w prosty i bezkosztowy sposób zaprogramować jednostkę sterującą.

Przeanalizowano ofertę najważniejszych producentów układów pod kątem dostępności w popularnych sklepach elektronicznych. Na podstawie tej analizy wybrano jednostkę opartą o procesor firmy Intel, kompatybilny z popularnym oprogramowaniem Arduino – Intel Curie Quark ARC (rysunek 4.2). Układ spełniał wszystkie wcześniej sprecyzowane założenia, a jego niewielkie rozmiary (23,5mm x 43,0mm) spowodowały, że rękawica nie utraciła ergonomii. Jednostka posiada wydajny procesor (32MHz, 196kB pamięci *flash*), zintegrowany 6-osiowy akcelerometr oraz moduł BLE (ang. *Bluetooth Low Energy*) w wersji 4.0 (Dfrobot, 2019).



Rys. 4.2. Intel Curie Quark ARC (Dfrobot, 2019)

Kolejnym kluczowym elementem rękawicy były analogowe czujniki ugięcia palców. Aby jak najdokładniej odwzorowywać pozycję zgięcia palców zdecydowano zastosować dwa rozmiary czujników: 5,5cm dla kciuka i małego palca oraz 7,3cm dla pozostałych trzech

palców. Wybrano rezystancyjne czujniki ugięcia firmy SparkFun (rysunek 4.3), które charakteryzują się liczbą cykli zgięcia na poziomie jednego miliona, a także temperaturą pracy w zakresie od  $-35^{\circ}\text{C}$  do  $+80^{\circ}\text{C}$  (Botland, 2019e).



Rys. 4.3. Czujniki ugięcia firmy SparkFun (Botland, 2019e)

#### 4.2.2. Schemat elektryczny

Dla układu elektronicznego urządzenia narysowano schemat elektryczny znajdujący się w załączniku 1 niniejszej pracy. Wiele połączeń wyeliminowano, ponieważ zastosowano układ Intel Curie Quark ARC, który posiada zintegrowany moduł komunikacyjny Bluetooth oraz akcelerometr z żyroskopem. Potencjał dodatni czujników ugięcia podłączony został do pinu 5V0 (tj. napięcie 5VDC) układu mikrokontrolera, a wszystkie wyprowadzenia sensorów połączono w konfiguracji szeregowej. Do wyprowadzeń analogowych (A0-A4) doprowadzono przewody sygnałowe, z każdego czujnika ugięcia, odpowiednio:

- A0 – kciuk,
- A1 – palec wskazujący,
- A2 – palec środkowy,
- A3 – palec serdeczny,
- A4 – palec najmniejszy.

Do każdego z wyżej wymienionych wyprowadzeń analogowych przyłączony został rezystor  $660\Omega$  doprowadzony do masy (tzw. GND). Zastosowanie rezystora podciągającego do masy ma na celu wyeliminowanie stanów nieustalonych na wejściach analogowych mikrokontrolera. Cały układ może być zasilany przez standardowy przewód ze złączem micro USB typu B lub przez bezpośrednie podpięcie potencjałów baterii do odpowiednich



wyprowadzeń płytki: BAT oraz GND. Wszystkie połączenia przewodowe zostały wykonane za pomocą elastycznych izolowanych przewodów miedzianych oraz cyny lutowniczej.

### 4.2.3. Software

Do stworzenia kompletnego oprogramowania dla urządzenia komunikacji głosowej niezbędna była znajomość dwóch bezpłatnych środowisk programistycznych Android Studio (Java, SQL) oraz Arduino IDE (język programowania C). Kluczem do prawidłowego działania całego urządzenia był odpowiedni podział zadań pomiędzy aplikację mobilną a mikrokontroler znajdujący się na rękawicy, co pokazane zostało na rysunku 4.4.

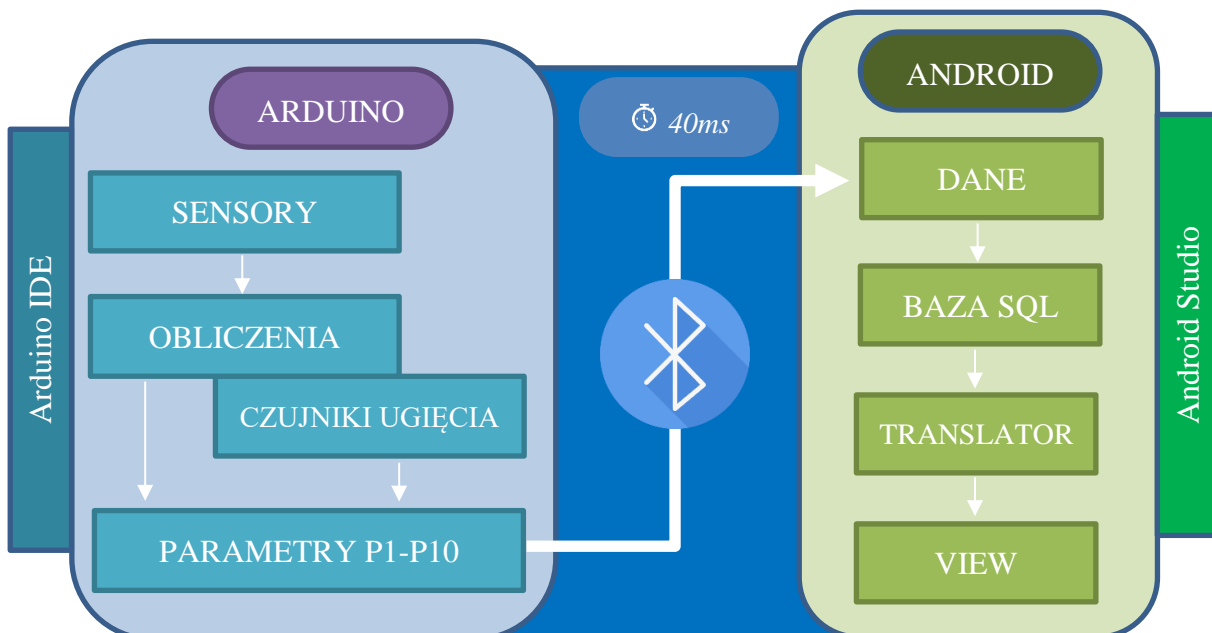
Android Studio pozwala tworzyć aplikacje mobilne na platformę Android, wykorzystując wszystkie dostępne w telefonach komórkowych interfejsy oraz urządzenia, np. GPS, Bluetooth, WiFi, aparat, itp. W stworzonej aplikacji zastosowana została komunikacja Bluetooth, lokalna baza danych SQLite, funkcjonalność dostępu do stron internetowych przez sieć komórkową, możliwość zapisu i odczytu danych z pliku oraz wbudowany syntezytor mowy. Struktura programu napisanego dla platformy Android opiera się na sekwencji wywoływania kilku funkcji:

1. Funkcja inicjalizacyjna. Przy uruchomieniu aplikacji ładują się zapisane dane użytkownika, takie jak nazwa ostatnio podłączonego urządzenia Bluetooth, parametry komunikacji bezprzewodowej, czas cyklu odświeżania danych oraz ustawienia syntezytora mowy i tłumacza. Funkcja inicjalizacyjna sprawdza czy aplikacja posiada niezbędne zgody na dostęp do lokalizacji oraz plików, a także inicjalizuje syntezytor mowy oraz interfejs Bluetooth – automatycznie nawiązuje połączenie z rękawicą.
2. Funkcja cykliczna. Wykonuje się co zadany okres czasu (standardowo 40ms), wyzwalane są wówczas odczyty parametrów przez Bluetooth. Odczytane dane są sprawdzane pod kątem niepowtarzalności z poprzednimi odczytami. Jeśli dane są świeże, zostają zarchiwizowane w tabeli dwuwymiarowej, a następnie wyzwalana jest funkcja obsługująca przeszukiwanie bazy SQLite. Gdy odnaleziona zostanie odpowiednio dopasowanie danych do słowa, aktywowany jest syntezytor mowy, funkcja tłumacząca oraz aktualizacja elementów *view* (tj. pól tekstowych).
3. Funkcja obsługi bazy danych SQLite. Jej zadaniem jest porównywanie danych odczytywanych w czasie rzeczywistym z danymi w bazie.

4. Funkcja obsługi komunikacji Bluetooth Low Energy. Odpowiada ona za odpytywanie urządzenia peryferyjnego o odpowiednie charakterystyki oraz zapisywanie odebranych danych do tablicy.
5. Funkcja tłumacząca tekst. Opiera się na wykorzystaniu sieci komórkowej. Tworzone jest odpowiednie zapytanie HTTP, zawierające ciąg znaków do przetłumaczenia oraz kierunek tłumaczenia, następnie jest ono wyzwalane. W informacji zwrotnej otrzymywany jest ciąg znaków będący tłumaczeniem.

Arduino IDE pozwala programować wiele popularnych układów mikroprocesorowych, wyposażonych w różnorodne urządzenia peryferyjne, np. czytniki RFID, moduł odcisku palca, Bluetooth i wiele innych. W aplikacji rękawicy wykorzystywana jest komunikacja Bluetooth, odczyty z wbudowanych sensorów oraz wejść analogowych.

Struktura programu w Arduino IDE składa się z części inicjalizacyjnej (tzw. *setup*) oraz pętli nieskończonej (*loop*). W funkcji *setup* deklarowane są podstawowe interfejsy dla Bluetooth oraz akcelerometru. W pętli głównej odczytywane są podstawowe parametry, tj. wartości zgięcia palców oraz odczyty z wbudowanych sensorów. Dalej wykonywane są obliczenia podstawowych parametrów charakterystycznych na podstawie aktualnych i historycznych danych przechowywanych w tablicach. Wartości ugięć oraz obliczone atrybuty opisowe zostają finalnie wstawione do tablicy zadeklarowanej charakterystyki, z której aplikacja mobilna odczytuje dane przez Bluetooth.



Rys. 4.4. Ogólny schemat kooperacji mikrokontrolera na rękawicy z aplikacją mobilną (opracowanie własne)

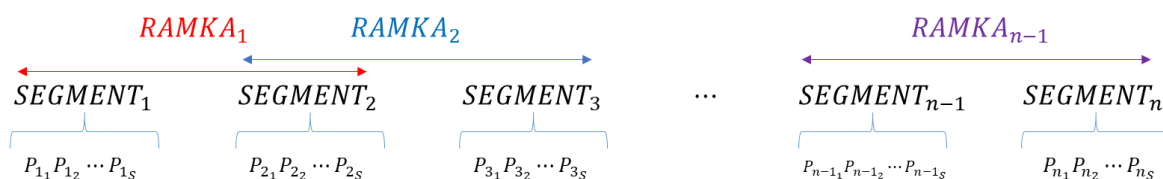
## 4.3. Budowa oprogramowania sterującego

### 4.3.1. Struktura gestu

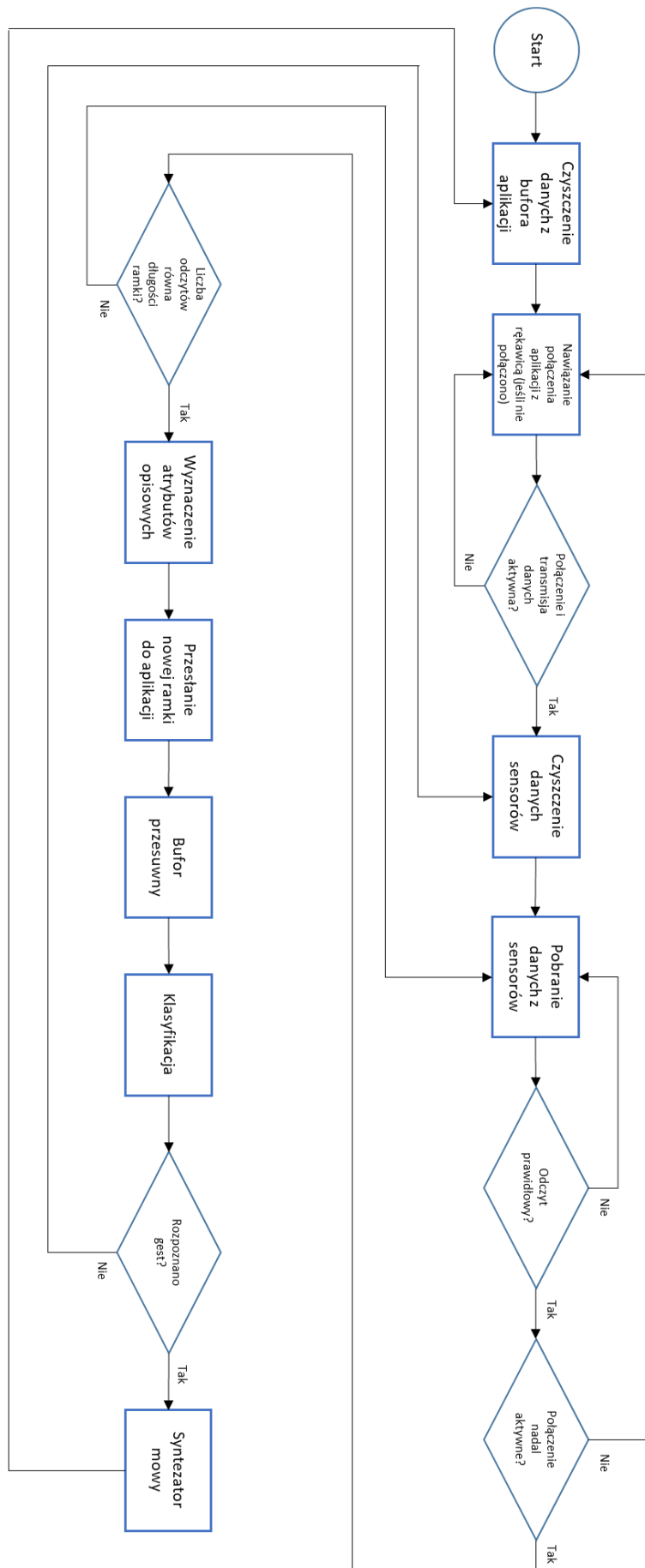
Oprogramowanie sterujące stanowi główny element obliczeniowo-decyzyjny w procesie klasyfikacji gestów wykonywanych przy pomocy rękawicy. Aby mieć możliwość identyfikowania wielu różnych ruchów, należy w pierwszej kolejności doprowadzić do unifikacji reprezentacji matematycznej gestu, aby każda pozycja w bazie danych (słowniku) oraz wykonywane w czasie rzeczywistym ruchy dłoni w przestrzeni były interpretowane w ten sam sposób. Jest to warunkiem koniecznym w celu umożliwienia porównania obu tych instancji.

Zaproponowany w pracy model matematyczny gestu składa się z szeregu atrybutów opisowych (wartości charakterystycznych), obliczonych na podstawie odczytu wartości sensorów. Uproszczony proces, przedstawiający ideę pobierania i obróbki danych w tej aplikacji, został przedstawiony na rysunku 4.6. Po nawiązaniu połączenia mikrokontrolera z aplikacją mobilną następuje rozpoczęcie pobierania danych z sensorów. Każdy odczyt jest weryfikowany pod kątem poprawności a następnie wpisywany jest do bufora. Gdy bufor osiągnie wymaganą długość, wówczas wykonywane są obliczenia, których wynikiem są kolejne atrybuty opisowe gestu. Po zakończeniu tego procesu, dane zostają przesłane do aplikacji mobilnej, gdzie następuje wpisanie ich w odpowiednie miejsce bufora wartości charakterystycznych. Po tym działaniu, rozpoczyna się proces klasyfikacji gestu. W przypadku, gdy jedno ze słów z bazy danych zostanie poprawnie przyporządkowane wykonanemu gestowi, wówczas trafia ono do syntezy mowy aplikacji, natomiast jeśli żadne ze słów nie zostały zaklasyfikowane do ruchu dłoni, rozpoczyna się ponowny proces pobierania danych z czujników.

Na rysunku 4.5 zaprezentowana została struktura zbiorów danych, z których złożony jest gest. Jak widać na grafice,  $n$  segmentów tworzy  $n-1$  ramek, które się wzajemnie zazębiają: każde dwie sąsiadujące ramki posiadają jeden wspólny segment danych. Segment z kolei składa się  $S$  próbek, z których każda jest pojedynczym odczytem z jednego z sensorów umieszczonych na rękawicy.



Rys. 4.5. Struktura danych w ramce komunikacyjnej (opracowanie własne)



Rys. 4.6. Struktura procesu pobierania i obróbki danych przez aplikację mobilną (opracowanie własne)

Uogólniony, matematyczny zapis każdego gestu znajdującego się w słowniku, można zapisać w postaci zaprezentowanej na rysunku 4.7. Skrótom  $AO$  oznaczono tutaj atrybut opisowy. Z poniższego schematu można zauważyć, że gest opisuje  $m$  atrybutów opisowych, z których każdy jest jednowymiarowym wektorem o długości  $n$ . W fazie testów, brane były pod uwagę różne atrybuty opisowe, jednak wyselekcjonowane zostały tylko te, które wносиły wartość dodaną w procesie klasyfikacji i maksymalizacji różnic między kolejnymi gestami. Długość wektorów została dobrana doświadczalnie, na bazie szeregu testów.

$$\begin{array}{cccccc}
 AO_{1_1} & AO_{1_2} & AO_{1_3} & \cdots & AO_{1_n} \\
 AO_{2_1} & AO_{2_2} & AO_{2_3} & \cdots & AO_{2_n} \\
 AO_{3_1} & AO_{3_2} & AO_{3_3} & \cdots & AO_{3_n} \\
 \vdots & \vdots & \vdots & & \vdots \\
 AO_{m_1} & AO_{m_2} & AO_{m_3} & \cdots & AO_{m_n}
 \end{array}$$

Rys. 4.7. Ogólna struktura rozkładu atrybutów opisowych (opracowanie własne)

Ostateczna forma struktury gestu, jaka została przyjęta w niniejszej pracy, została zaprezentowana na rysunku 4.8. Rozgraniczyć można tutaj dwie główne części: macierz cech charakterystycznych (oznaczona na grafice literą  $C$ ) oraz macierz uśrednionych wartości ugięcia każdego z pięciu palców dłoni (oznaczone na rysunku literą  $P$ ). Długość wektora została ustalona na poziomie równym 9. Każdy z gestów składa się zatem łącznie z 90 liczb, które są bezpośrednio przetwarzane przez opracowany klasyfikator.

$$\begin{array}{cccccccc}
 C_{1_1} & C_{1_2} & C_{1_3} & C_{1_4} & C_{1_5} & C_{1_6} & C_{1_7} & C_{1_8} & C_{1_9} \\
 C_{2_1} & C_{2_2} & C_{2_3} & C_{2_4} & C_{2_5} & C_{2_6} & C_{2_7} & C_{2_8} & C_{2_9} \\
 C_{3_1} & C_{3_2} & C_{3_3} & C_{3_4} & C_{3_5} & C_{3_6} & C_{3_7} & C_{3_8} & C_{3_9} \\
 C_{4_1} & C_{4_2} & C_{4_3} & C_{4_4} & C_{4_5} & C_{4_6} & C_{4_7} & C_{4_8} & C_{4_9} \\
 C_{5_1} & C_{5_2} & C_{5_3} & C_{5_4} & C_{5_5} & C_{5_6} & C_{5_7} & C_{5_8} & C_{5_9} \\
 P_{1_1} & P_{1_2} & P_{1_3} & P_{1_4} & P_{1_5} & P_{1_6} & P_{1_7} & P_{1_8} & P_{1_9} \\
 P_{2_1} & P_{2_2} & P_{2_3} & P_{2_4} & P_{2_5} & P_{2_6} & P_{2_7} & P_{2_8} & P_{2_9} \\
 P_{3_1} & P_{3_2} & P_{3_3} & P_{3_4} & P_{3_5} & P_{3_6} & P_{3_7} & P_{3_8} & P_{3_9} \\
 P_{4_1} & P_{4_2} & P_{4_3} & P_{4_4} & P_{4_5} & P_{4_6} & P_{4_7} & P_{4_8} & P_{4_9} \\
 P_{5_1} & P_{5_2} & P_{5_3} & P_{5_4} & P_{5_5} & P_{5_6} & P_{5_7} & P_{5_8} & P_{5_9}
 \end{array}$$

Rys. 4.8. Szczegółowa struktura atrybutów opisowych w ramach komunikacyjnych (opracowanie własne)

### 4.3.2. Baza słów środowiska produkcyjnego

W słowniku prototypowej aplikacji znajduje się 50 słów, jak przedstawiono w poniższym zestawieniu (tabela 4.2):

Tab. 4.2. Zestawienie słów i wyrażen słownika aplikacji mobilnej (opracowanie własne)

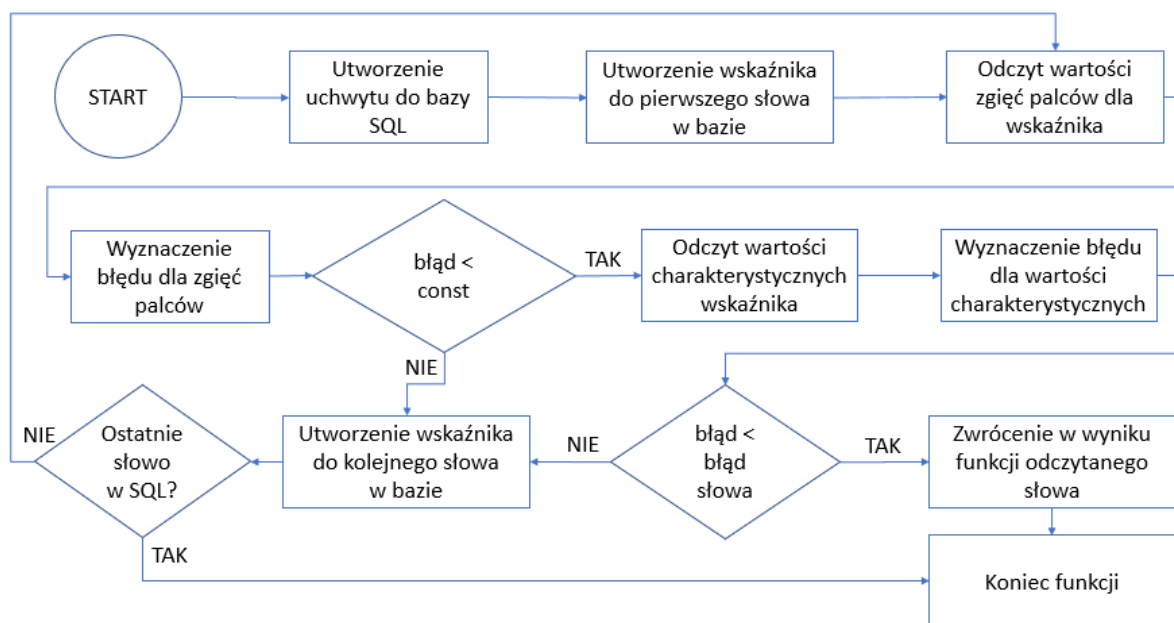
magister	technologia	nazywam się	produkować	efekt
tytuł	rękawice	jak masz na imię	przestrzeń	inżynier
przemysł	wskaźnik	miło cię poznać	rysunek	pamięć
automatyczny	pracownik	migać	siła	równoległy
czas	cześć	jeszcze raz	wykres	część
cykl	niedosłyszający	chcieć	wartość	próbka
robot	głuchy	otrzymać	wypadek	stan
system	tak	koszt	natężenie	maszyna
analizować	instalacja	pomoc	napięcie	porównanie
asystować	jakość	praca	dokument	skok

Stworzona lokalnie baza danych SQL składa się z 95 kolumn z czego 90 z nich odpowiada za parametry charakterystyczne, natomiast pozostałe to: słowo lub wyrażenie, dopuszczalny błąd dopasowania, rozróżnienie: części mowy (rzeczownik, czasownik, przymiotnik, itd.), rodzaju (męski, żeński, nijaki) oraz liczby (pojedynczej lub mnogiej). Trzy ostatnie wymienione parametry będą służyły do prawidłowego formułowania zdań, gdyż języki migowe, tak jak języki narodowe posiadają swoją własną gramatykę. W poniższym spisie (tabela 4.3) pokazano strukturę tabeli w bazie SQL.

Tab. 4.3. Wizualizacja danych w tabeli bazy SQL (opracowanie własne)

słowo	błąd	część mowy	rodzaj	liczba	P1111	P1112	P1113	... (86parametrów)...	P1920
cześć	23	rzeczownik	brak	brak	1567	3662	904	...	70
niedosłyszający	34	rzeczownik	brak	pojedyncza	953	3736	332	...	87
chcieć	31	czasownik	brak	brak	1321	2345	553	...	66
...	...	...	...	...	...	...	...	...	...

Sposób wyszukiwania słów i wyrażen w bazie danych jest realizowany dwuetapowo, i został przedstawiony na rysunku 4.9. Analizowane są wszystkie słowa w bazie. W pierwszym etapie odczytywane są wartości zgięć palców, czyli 45 liczb. Wartości te są porównywane do odczytów rzeczywistych, po czym wyznaczany jest błąd, pomiędzy każdą z wartości. Jeśli błąd nie przekracza wcześniej zadeklarowanej stałej wartości, to algorytm przechodzi do drugiego etapu, czyli porównywania parametrów charakterystycznych. Ponownie, przy użyciu opracowanego klasyfikatora wyznaczany jest błąd pomiędzy wartościami z bazy SQL a wartościami aktualnymi, odebranymi przez Bluetooth. Jeśli wyznaczony błąd jest niższy niż wartość zadeklarowana w bazie, to słowo zostaje zwrócone jako wynik działania funkcji. W przypadku szczególnym, gdzie kilka słów spełnia wszystkie warunki, wybierane jest to, które ma najniższy błąd.



Rys. 4.9. Schemat blokowy dla funkcji przeszukującej bazę słów (opracowanie własne)

Gesty w językach migowych można podzielić ze względu na dynamikę zmiany niektórych parametrów:

1. Statyczny układ palców oraz dynamiczna zmiana ruchu ręki. Takie gesty jest stosunkowo łatwo rozpoznać, ponieważ zmienna jest tylko jedna grupa wartości (są to przeliczone wartości charakterystyczne na podstawie odczytów z sensorów).
2. Dynamiczna zmiana układu palców oraz brak ruchu ręki. Ten rodzaj gestów wykorzystuje się do pokazania poszczególnych liter alfabetu. Z uwagi na fakt korzystania z czujników ugięcia palców, o dużej rozdzielczości rozpoznanie poszczególnych liter nie stanowi większego problemu.
3. Dynamiczna zmiana palców oraz ruchu ręki. Jest to trudniejsza grupa gestów w kontekście rozpoznania, ponieważ wymaga stosunkowo dokładnego odwzorowania danego gestu w odpowiednim czasie. Ważne, aby zmiana ułożenia palców następowała za każdym razem, w tym samym momencie ruchu ręki.

#### 4.4. Metodyka testowania rozwiązania

Opisywane w niniejszej pracy urządzenie do komunikacji głosowej dla osób głuchoniemych w warunkach produkcyjnych poddane zostało wielokrotnym testom na każdym etapie zaawansowania prac. Najważniejsze jednak wyniki w kontekście powodzenia całego

projektu płyną z testów wykonywanych na gotowym urządzeniu z zaimplementowanymi w finalnej wersji algorytmami klasyfikacji gestów.

W tabeli 4.4 przedstawiono przykładowy schemat podziału próbek przy wykonywaniu testu na zbiorze przykładów należących do danej klasy. Wartość *PP* (prawdziwie pozytywne, ang. *true positive*) oznacza zbiór próbek, które należąc do klasy *A* zostały również do niej zaklasyfikowane, *FN* (fałszywie negatywne, ang. *false negative*) określa nieprawidłowe wskazanie klasy innej niż *A* (zamiast klasy *A*), *FP* (fałszywie pozytywne, ang. *false positive*) reprezentuje zbiór przykładów z nieprawidłowym wskazaniem na wyróżnioną klasę *A*, natomiast *PN* (prawdziwie negatywne, ang. *true negative*) oznacza prawidłowe wskazanie na jedną z pozostałych klas (inną niż klasa *A*).

Tab. 4.4. Schematyczny zapis rozkładu próbek przy wyznaczaniu współczynników czułości i specyficzności (opracowanie własne na podstawie (Chmielnicki, 2012, str. 38.39))

	Gest należy do klasy A	Gest należy do innej klasy niż A
Przydzielenie gestu do klasy A	PP	FP
Przydzielenie gestu do klasy innej niż A	FN	PN

Aby kompleksowo określić jakość wykonanego prototypu rękawicy, została ona przetestowana w szerokim spektrum funkcjonalności. Określone zostały kolejno:

1. Wytrzymałość mechaniczna. Pod koniec wszystkich testów rękawica została dokładnie zbadana pod kątem występowania uszkodzeń mechanicznych, a także sprawdzone zostały połączenia elektryczne i mocowania podzespołów.
2. Średni czas wykonywania obliczeń dla ekstrakcji cech charakterystycznych. Jest to czas, w jakim wyznaczane są wartości atrybutów opisowych na mikrokontrolerze, a podczas którego nie są pobierane próbki danych z sensorów. Jest to zatem okres, z którego dane o ruchu rękawicy i poziomie ugięcia palców są bezpowrotnie tracone. Określony został przy tym stopień istotności takiego ubytku.
3. Średni czas procesu klasyfikacji. Jest to okres pomiędzy pobraniem nowych danych przez aplikację a zakończeniem działań klasyfikatora. W przypadku gdy czas ten osiągałby zbyt duże wartości, wówczas mogłoby dojść do utraty ramek wysyłanych przez mikrokontroler lub do zmian w stabilności i płynności komunikacji



bezprzewodowej. Sprawdzony został również wpływ liczby słów w słowniku na wzrost średniego czasu procesu klasyfikacji.

4. Zapas błędu. Jest to procent jaki stanowi błąd średni w błędzie granicznym. Każdy z wykonywanych gestów obarczony jest pewną wielkością błędu. Indywidualnie, dla wszystkich instancji w bazie słów określona jest wartość błędu granicznego, czyli maksymalnego akceptowalnego. Błąd średni danego słowa stanowi wartość średnią wszystkich błędów, które zostały zarejestrowane w fazie testowania przy wykonywaniu tego gestu.
5. Czułość. Stanowi ona iloraz wszystkich poprawnych klasyfikacji gestów dla testowanej klasy A i łącznej sumy wykonanych gestów w tym teście. Współczynnik czułości oblicza się według wzoru 4.1 (Chmielnicki, 2012, str. 38-39).

$$S_n = \frac{PP}{PP + FN} \cdot 100\% [\%] \quad (4.1)$$

gdzie:

$S_n$  – współczynnik czułości,

$PP$  – suma wyników prawdziwie pozytywnych (patrz tabela 4.4),

$FN$  – suma wyników fałszywie pozytywnych (patrz tabela 4.4).

6. Specyficzność. Zdefiniowana jest jako iloraz wszystkich poprawnych klasyfikacji gestu dla klasy innej niż klasa A i łącznej sumy wszystkich wykonanych gestów w tym teście. Współczynnik specyficzności oblicza się według wzoru 4.2 (Chmielnicki, 2012, str. 38-39).

$$S_p = \frac{PN}{PN + FP} \cdot 100\% [\%] \quad (4.2)$$

gdzie:

$S_p$  – współczynnik specyficzności,

$PN$  – suma wyników prawdziwie negatywnych (patrz tabela 4.4),

$FP$  – suma wyników fałszywie pozytywnych (patrz tabela 4.4).

7. Skuteczność. Określona została w kontekście całościowym oraz indywidualnym. W tak interpretowanej macierzy klasyfikacji jak w tabeli 4.4, skuteczność można przedstawić za pomocą dwóch wzorów. W pierwszym z nich stanowi ona iloraz wszystkich poprawnych przyporządkowań dla klasy A i innej niż A do łącznej sumy wszystkich gestów w tych dwóch testach (wzór 4.3 – skuteczność indywidualna). W drugim przypadku skuteczność rozpatrywana jest jako iloraz sumy wszystkich

poprawnych klasyfikacji i sumy wszystkich wykonanych gestów (wzór 4.4 – skuteczność całościowa) (Chmielnicki, 2012, str. 38-39).

$$Acc_i = \frac{PP + PN}{PP + PN + FP + FN} \cdot 100\% [\%] \quad (4.3)$$

gdzie:

$Acc_i$  – współczynnik skuteczności indywidualnej,

$PP$  – suma wyników prawdziwie pozytywnych (patrz tabela 4.4),

$FN$  – suma wyników fałszywie pozytywnych (patrz tabela 4.4),

$PN$  – suma wyników prawdziwie negatywnych (patrz tabela 4.4),

$FP$  – suma wyników fałszywie pozytywnych (patrz tabela 4.4).

$$Acc_c = \frac{c_1 + c_2 + c_3 + \dots + c_N}{n_1 + n_2 + n_3 + \dots + n_N} \cdot 100\% = \frac{C}{N} \cdot 100\% [\%] \quad (4.4)$$

gdzie:

$Acc_c$  – współczynnik skuteczności całościowej,

$C = c_1 + c_2 + c_3 + \dots + c_N$  – liczba wszystkich poprawnie rozpoznanych próbek,

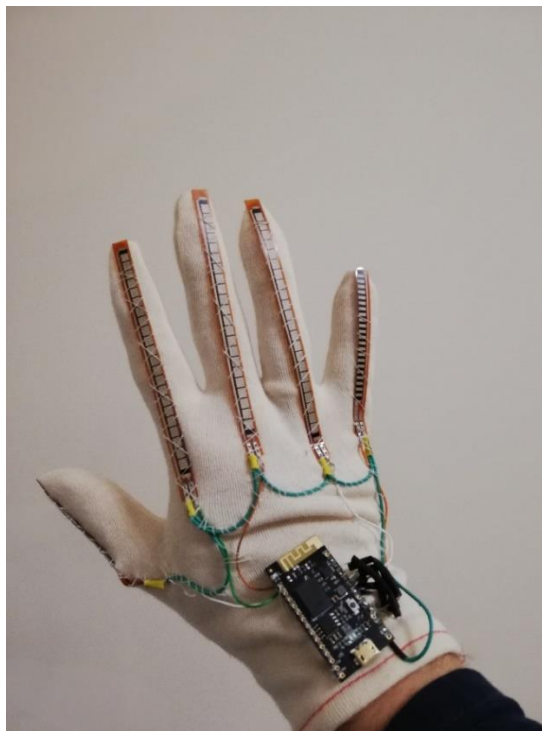
$N = n_1 + n_2 + n_3 + \dots + n_N$  – liczba próbek w zbiorach testowych.

Wszystkie powyższe właściwości zostały obliczone na bazie zbiorów testowych, w których skład wchodzi zarówno dane przygotowane przed procesem uczenia, do weryfikacji działania klasyfikatora oraz dane zebrane podczas testów po działaniach optymalizacyjnych procesu klasyfikacji (Chmielnicki, 2012, str. 38, 39).

## 5. Wyniki pracy

### 5.1. Uzyskany prototyp rozwiązania

Podczas działań nad niniejszą pracą dyplomową skonstruowany został prototyp funkcjonalny rękawicy wraz z aplikacją mobilną dla urządzeń opartych o platformę Android. Wszystkie elementy elektroniki zostały przymocowane na bawełnianej, elastycznej i antystatycznej rękawicy, jak pokazano na rysunku 5.1. Całość była testowana wielokrotnie przy kolejnych fazach rozwoju aplikacji i nie napotkano większych problemów związanych z odrywaniem się elementów elektronicznych od rękawicy lub niestabilnej pracy elektroniki, poza jednym przypadkiem, w którym jeden z przewodów czujnika ugięcia uległ naderwaniu. Urządzenie nie zostało wyposażone jak dotąd we własną baterię, dlatego jest zasilanie przez kabel USB podłączony do zewnętrznej baterii (typu *powerbank*).

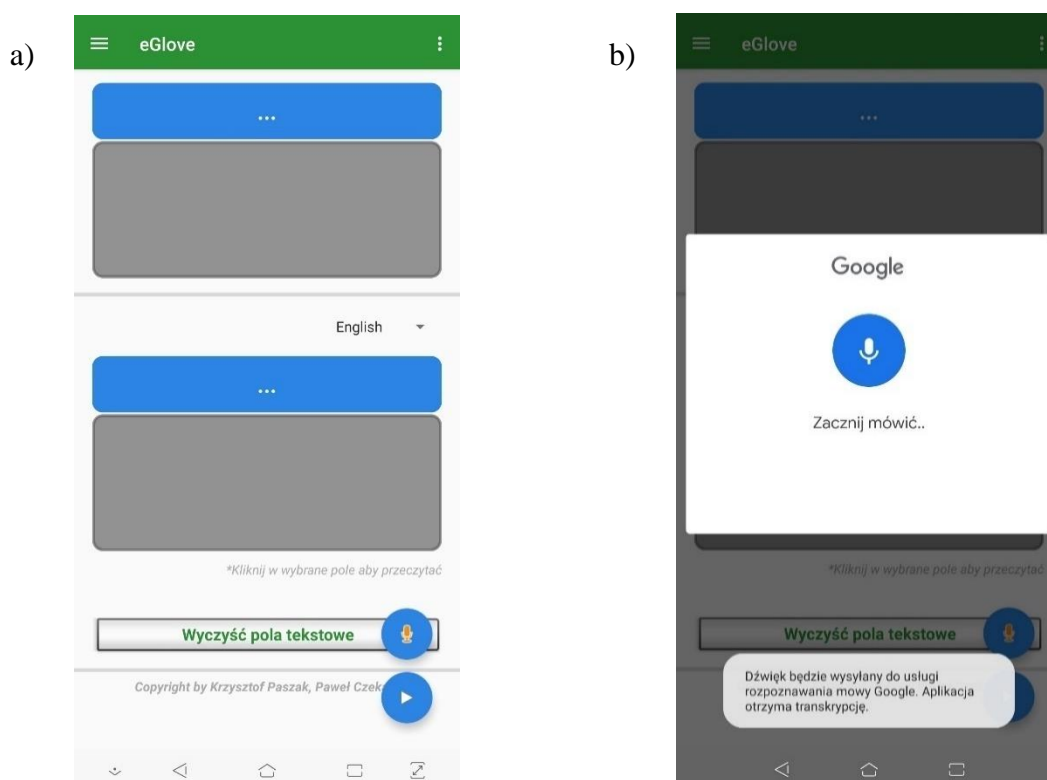


Rys. 5.1. Prototyp rękawicy urządzenia komunikacji głosowej (opracowanie własne)

Rękawica komunikuje się bezprzewodowo z urządzeniem mobilnym. Przetestowany zasięg dla poprawnej komunikacji oscylował w granicach 10m w otwartej przestrzeni. W przypadku normalnego trybu pracy, tj. odległość ok. 2m między rękawicą a telefonem komórkowym oraz nagłych, szybkich ruchów ręki, nie napotkano na przypadki zerwania połączenia Bluetooth.

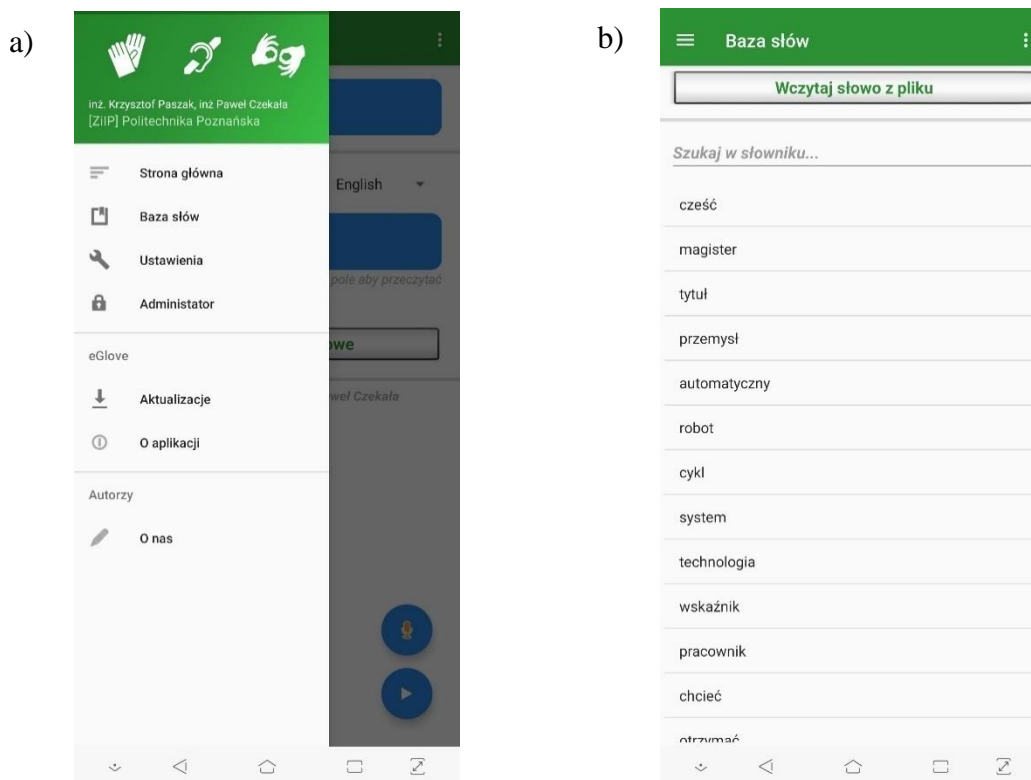
Aplikacja mobilna została dostosowana dla urządzeń z systemem Android w wersji 5.1 lub wyższej. Jej struktura i układ zostały zaprojektowane w taki sposób, aby obsługa była intuicyjna. Ekran startowy składa się z dwóch części oraz dwóch przycisków, co pokazane

zostało na rysunku 5.2a. W pierwszym, niebieskim polu tekstowym wyświetlany jest aktualnie odczytany gest w postaci słowa lub wyrażenia w języku polskim. W kolejnym polu dodawane są wszystkie odczytane wyrażenia, jest więc dostęp do historii odczytanych gestów. Bliźniacze pola poniżej ukazują te same informacje, ale w języku obcym, zdefiniowanym przez użytkownika. Tłumaczenia można wyłączyć przechodząc do ustawień. Przycisk w prawym dolnym rogu, oznaczony ikoną białego trójkąta, służy do aktywacji funkcji rozpoznawania gestów. Natomiast przycisk, z pomarańczową ikoną mikrofonu, włącza systemową funkcję konwersji mowy na tekst (rysunek 5.2b). Przeciągając palcem od lewej krawędzi ekranu, można otworzyć panel zakładek, jak na rysunku 5.3a.



Rys. 5.2 Aplikacja mobilna a. Ekran główny b. Ekran rozpoznawania mowy (opracowanie własne)

Z poziomu menu głównego można przejść do zakładki „Baza słów”, w której znajdują się wszystkie zdefiniowane w bazie SQL słowa (rysunek 5.3b). Niezalogowany użytkownik może jedynie podglądać słowa i dodawać je do bazy za pomocą funkcji wczytania z odpowiedniego pliku tekstowego. Zalogowany administrator może dodatkowo edytować poszczególne wartości charakterystyczne dla każdego słowa, co pokazano na rysunku 5.4a. Słownik aplikacji posiada funkcję wyszukiwania słów na podstawie wpisanej frazy.

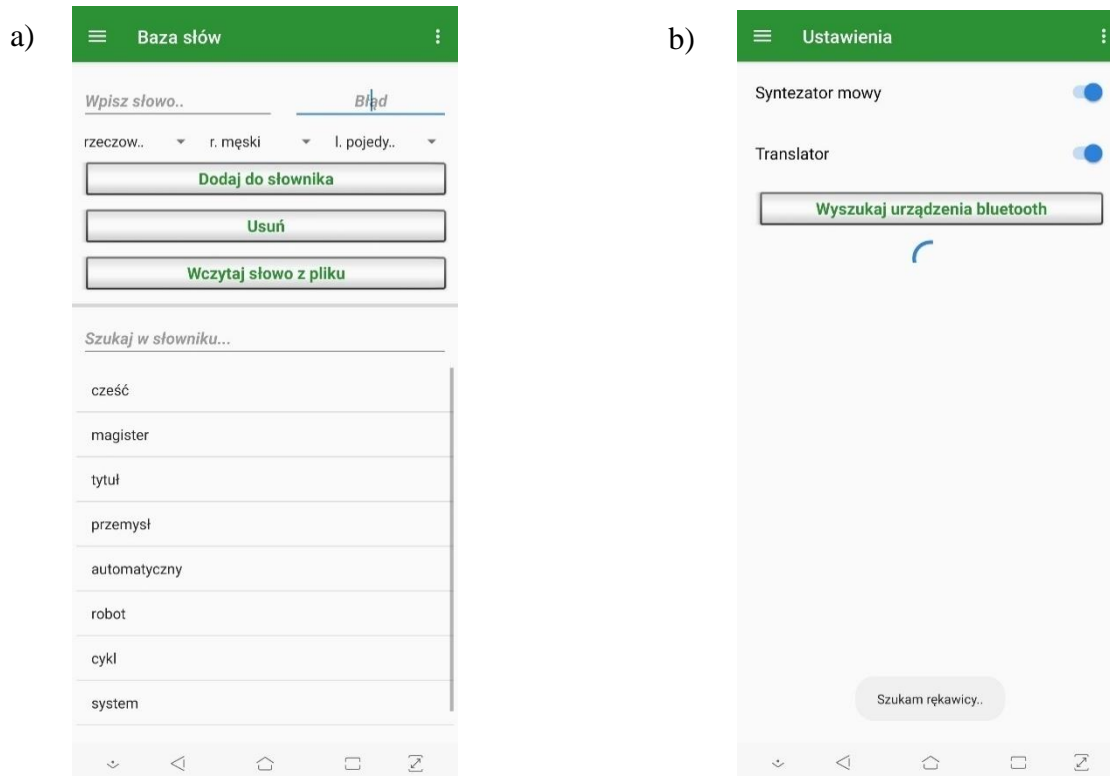


Rys.5.3 Aplikacja mobilna a. Ekran menu b. Baza słów w aplikacji (opracowanie własne)

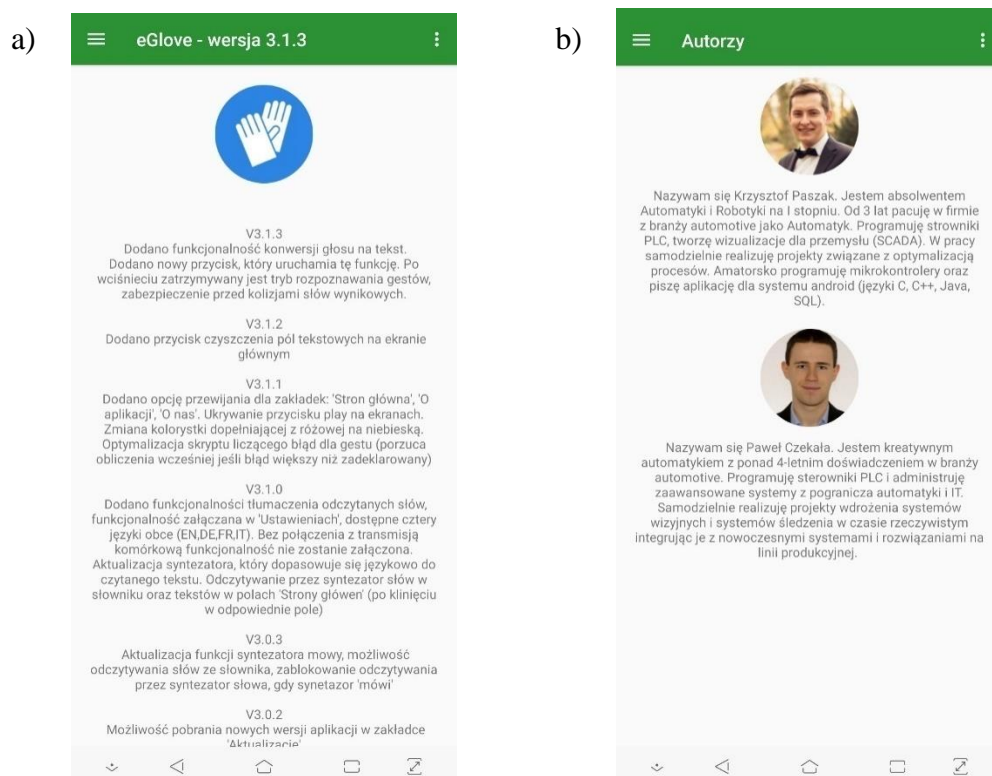
Na ekranie ustawień, pokazanym na rysunku 5.4b, użytkownik może dezaktywować lub aktywować syntezytor mowy oraz tłumaczenie na obcy język. Inną ważną funkcją jest wyszukiwanie urządzeń Bluetooth. Przy pierwszym uruchomieniu wymagane jest wyszukanie rękawicy oraz manualne połączenie się z nim poprzez kliknięcie na nazwę urządzenia na liście. Przy kolejnych uruchomieniach aplikacji, ustanowienie połączenia będzie realizowane automatycznie.

Klikając w prawym górnym rogu aplikacji można przejść do okna logowania. Zalogowanie się jako administrator daje możliwość podglądu parametrów komunikacji w czasie rzeczywistym oraz umożliwia zmianę parametrów niektórych funkcji.

Aplikacja została również wyposażona w ekrany informacyjne. W zakładce „O aplikacji” znajdują się informacje na temat zmian w kolejnych wersjach software’u (rysunek 5.5a). Klikając na przycisk „O nas” można przeczytać podstawowe informacje o doświadczeniu zawodowym konstruktorów urządzenia (rysunek 5.5b).



Rys.5.4 Aplikacja mobilna a. Słownik administratora b. Ekran ustawień (opracowanie własne)

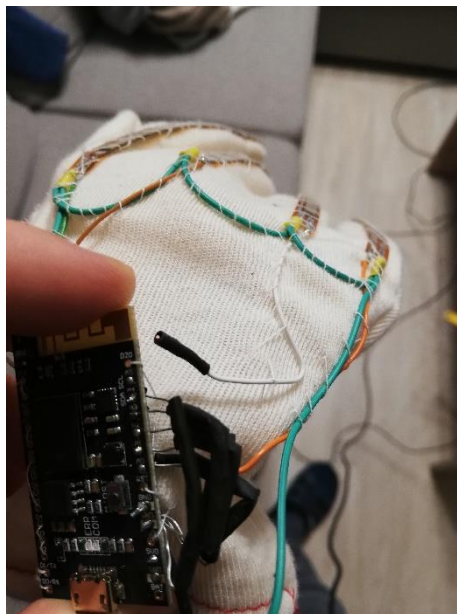


Rys. 5.5 Aplikacja mobilna a. Ekran „O aplikacji” b. Ekran o autorach (opracowanie własne)

## 5.2. Testy i ocena klasyfikacji

Ten podrozdział poświęcony został określeniu parametrów z jakimi prototyp rękawicy rozpoznaje gesty użytkownika. Ocenie zostały poddane aspekty opisane w podrozdziale 4.4. Testy przeprowadzone zostały na całej bazie słów, w dwóch trybach: *offline* i *online*. Testy *offline* polegały w głównej mierze na ocenie skuteczności klasyfikatora i jego optymalizacji na podstawie otrzymanych wyników. Bazę danych stanowiły w tym przypadku dane zgromadzone przed opracowaniem klasyfikatora, spośród których część została wykorzystana także do jego utworzenia. Testy *online* polegały już tylko na ocenie wszystkich parametrów jakościowych rękawicy na bazie gestów wykonywanych przez użytkownika na ostatecznej wersji klasyfikatora. Klasyfikacja następowała w tym przypadku w czasie rzeczywistym podczas wykonywania ruchów dłonią.

Pierwszym aspektem poddanym ocenie była wytrzymałość mechaniczna prototypu. Po przeprowadzeniu wszystkich testów nie zauważono żadnych uszkodzeń mechanicznych rękawicy ani sensorów. Materiał nie był naruszony w miejscach łączenia sensorów i mikrokontrolera z rękawicą. W przypadku połączeń elektrycznych stwierdzono jedno uszkodzenie, które znajdowało się w miejscu łączenia przewodu wychodzącego z czujnika ugięcia i wchodzącego na wejście analogowe mikrokontrolera. Po ręcznym poruszaniu tego przewodu połączenie między tymi dwoma elementami zostało ostatecznie całkowicie zerwane, przez co musiało zostać wykonane nowe połączenie. Na rysunku 5.6 pokazano omawiane uszkodzenie.



Rys. 5.6. Uszkodzenie połączenia elektrycznego rękawicy (opracowanie własne)

Na rysunku 5.7 zaprezentowany został rozkład czasów obliczeń ekstrakcji atrybutów opisowych przez mikrokontroler oraz czas klasyfikacji gestu przez aplikację mobilną. Kolejne czasy zostały zmierzone po dodaniu nowego słowa do słownika (zaznaczone na osi poziomej) i zostały uśrednione na podstawie kilkunastu prób dla każdego punktu pomiarowego. Jak widać na poniższym wykresie, średni czas wyznaczania atrybutów opisowych jest względnie stały na przestrzeni 50 wartości i waha się w zakresie od 16 ms do 22 ms. Przyjmując, że jedna ramka gestu trwa średnio 130 ms, wówczas czas obliczeń stanowi od 12% do 17% całej ramki.

Średni czas procesu klasyfikacji gestu rośnie w miarę zwiększania się bazy słów. Przy pomocy narzędzi arkusza kalkulacyjnego Excel wyznaczona została linia regresji dla wszystkich 50 punktów pomiarowych, która została opisana wzorem 5.1.

$$y = 0,3393x + 19,722 \quad (5.1)$$

gdzie:

$y$  – czas wyrażony w *ms*,

$x$  – liczba słów w słowniku.

Na tej podstawie można oszacować jaką maksymalną bazę słów może zawierać aplikacja, aby utrzymać jej konieczny warunek działania: rozpoznawanie gestów w trybie rzeczywistym, nie tracąc przy tym próbek danych wysyłanych przez mikrokontroler. Przyjmując, że gest trwa średnio 1,2 sekundy, można założyć, że aplikacja posiada właśnie tyle czasu na przeprowadzenia procesu klasyfikacji, gdyż pozostałe operacje związane z wyświetlaniem wyników i wykorzystanie syntezy mowy działają na osobnym wątku procesora telefonu. Maksymalna liczba słów w słowniku przy obecnej formie klasyfikatora będzie wówczas wynosić 3478 (wzór 5.2).

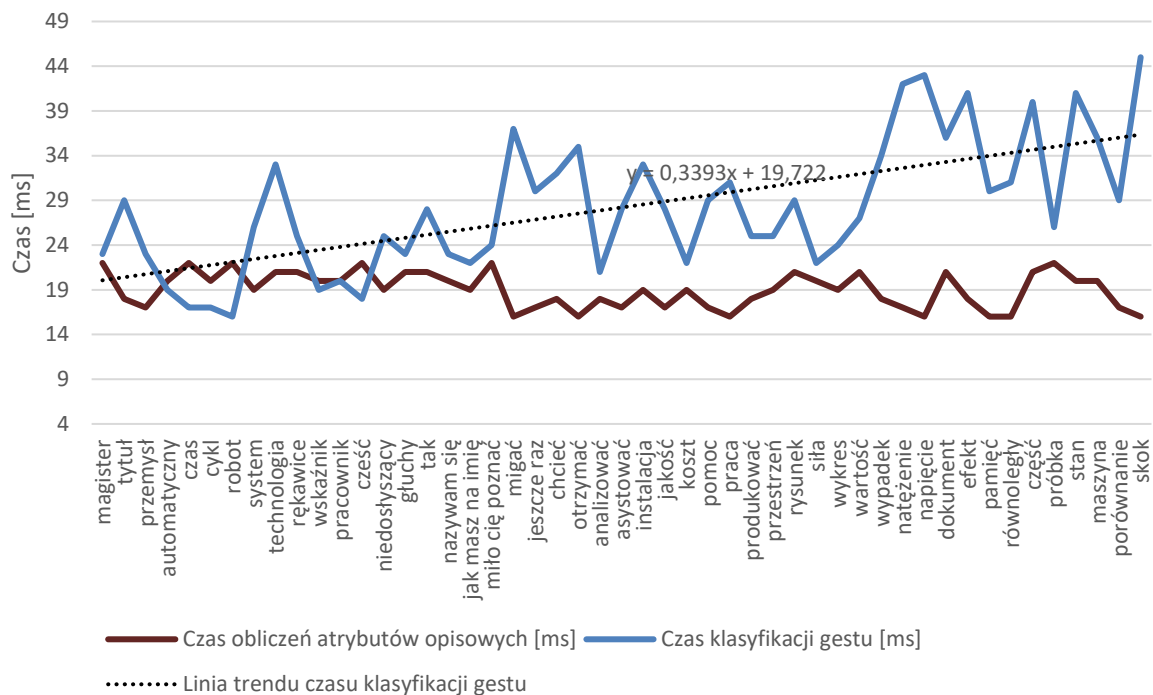
$$S_{max} = \frac{t_{gestu} - 19,722}{0,3393} = \frac{1200 - 19,722}{0,3393} \cong 3478 \text{ słów} \quad (5.2)$$

gdzie:

$S_{max}$  – maksymalna liczba słów w słowniku,

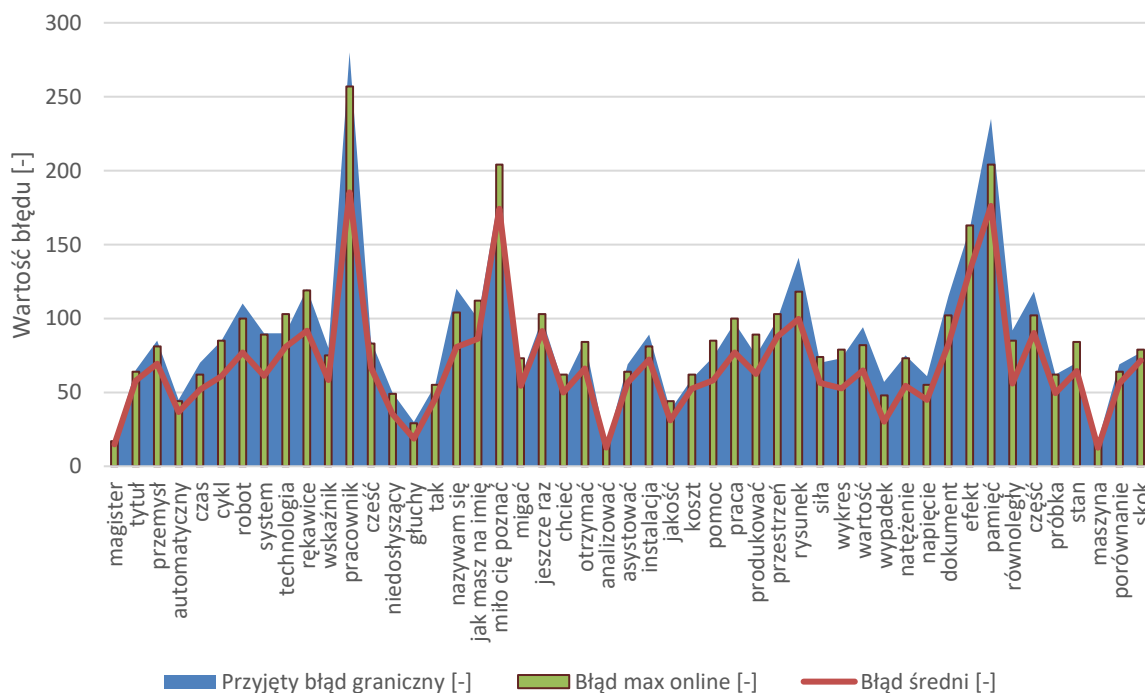
$t_{gestu}$  – średni czas trwania gestu wyrażony w *ms*.





Rys. 5.7. Rozkład czasów obliczeń atrybutów opisowych i klasyfikacji gestu dla rosnącej liczby gestów (opracowanie własne)

Rysunek 5.8 przedstawia rozkład błędów podczas procesu klasyfikacji konkretnych słów oznaczonych na osi poziomej. Przyjęty błąd graniczny oznacza jaki dla danego słowa został przyporządkowany maksymalny błąd w trakcie procesu klasyfikacji. Błąd max online określa maksymalny błąd podczas testów klasyfikacji każdego ze słów na przestrzeni 20 prób.



Rys. 5.8. Rozkład błędów dla dostępnych gestów w słowniku (opracowanie własne)

Błąd średni stanowi średnią z 20 prób wykonania każdego z gestów. Dla powyższego rozkładu jedynie w przypadku słów *analizować* oraz *maszyna* średnie wartości błędów prawie pokryły się z wartościami ustalonych błędów granicznych. Pozostałe słowa pomimo przekroczeń w kilku przypadkach wartości błędów granicznych, posiadają wartości średnie poniżej dopuszczalnych kryteriów.

Współczynniki czułości i specyficzności uzyskały średnie wyniki na poziomie odpowiednio 91% i 93%. Jednak z uwagi na to, że średnie rezultaty tych czynników nie są miarodajne w kontekście działania całego systemu, poniżej zaprezentowano tabelę 5.1, utworzoną na podstawie tej z podrozdziału 4.4 dla słowa *stan*, dla którego otrzymane zostały najgorsze rezultaty spośród badanych 50 wyrażzeń. Klasą A oznaczono wspomniany wyżej gest *stan*.

Tab. 5.1. Rozkład próbek dla klasy „stan”

	Gest należy do klasy A	Gest należy do innej klasy niż A
Przydzielenie gestu do klasy A	14	3
Przydzielenie gestu do klasy innej niż A	6	17

Na podstawie powyższej tabeli wyznaczone zostały parametry czułości i specyficzności, które stanowią graniczne wartości dla wszystkich badanych przykładów:  $S_n = 70\%$  i  $S_p = 85\%$ .

Współczynniki skuteczności zostały rozpatrzone w dwóch przypadkach: indywidualnym (dla każdego gestu) oraz globalnym (skuteczność całościowa). Skuteczność indywidualna została obliczona według wzoru 5.3, dla danych z tabeli 5.1, ponieważ dla gestu *stan* osiąga ona najgorszy rezultat.

$$Acc_i = \frac{14 + 17}{14 + 6 + 17 + 3} \cdot 100\% = 77,5\% \quad (5.3)$$

W poniższej tabeli 5.2 zgromadzono średnie wyniki z testów przeprowadzonych na 50 gestach, po 20 prób na każdy z gestów. Poprawne odczyty oznaczają prawidłowe przyporządkowanie gestu do klasy w momencie jego wykonania, błędne odczyty określają liczbę niezaklasyfikowanych gestów do danej klasy w momencie wykonania gestu (np. gdy wartość błędu przekroczyła przyjęty błąd graniczny) natomiast przypadkowe odczyty oznaczają zaklasyfikowanie gestu do danej klasy w sytuacji, gdy nie był wykonywany żaden

konkretny gest lub jakikolwiek inny zdefiniowany w słowniku. W tabeli podane są wartości średnie z testów.

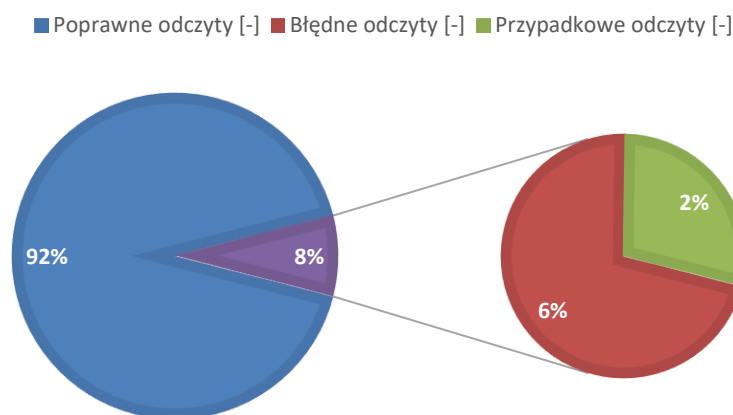
Tab. 5.2. Rozkład wszystkich odczytów gestów (opracowanie własne)

Poprawne odczyty [-]	Błędne odczyty [-]	Przypadkowe odczyty [-]
18,84	1,16	0,47

Ostatecznie zatem skuteczność całościowa wynosić będzie 92%, co obliczono na podstawie wzoru 5.4.

$$Acc_c = \frac{18,84}{18,84 + 1,16 + 0,47} \cdot 100\% \cong 92\% \quad (5.4)$$

Struktura poprawnych i niepoprawnych odczytów została przedstawiona na rysunku 5.9.



Rys. 5.9. Udział błędnych odczytów gestów w całej próbie (opracowanie własne)

### 5.3. Wyniki ankiety i testów grupy docelowej

W niniejszej pracy dyplomowej wykorzystano ankietę jako jedno z narzędzi do weryfikacji postępów pracy nad urządzeniem do komunikacji głosowej dla osób głuchoniemych w warunkach produkcyjnych. Badanie zostało przeprowadzone zarówno na grupie osób słyszących jak i wśród osób z niepełnosprawnościami słuchowymi, aby poznać szersze spektrum opinii na temat prototypu. W pierwszej części tego podrozdziału przedstawiona została struktura badanych osób, w drugiej zaprezentowano między innymi obraz problemów z komunikacją w życiu zawodowym osób niesłyszących i niedosłyszących z uwzględnieniem oceny wagi poszczególnych trudności przy użytkowaniu rękawic. Trzecia sekcja porusza te same kwestie co poprzednia, natomiast grupą docelową były osoby słyszące lub nieznaną

języka migowego. W pracy nie zostały poddane szczegółowej analizie wypowiedzi osób nieaktywnych zawodowo (obecnie lub w przeszłości), ponieważ osoby te nie stanowią głównej grupy docelowej dla zaprojektowanego prototypu inteligentnych rękawic.

Kwestionariusz ankiety został stworzony przy pomocy narzędzi Formularzy Google ([www.google.pl/intl/pl/forms/about/](http://www.google.pl/intl/pl/forms/about/)). Badanie zostało przeprowadzone we wrześniu 2019 roku. W ankiecie łącznie wzięło udział 210 osób, z czego 112 (53%) stanowiły kobiety a 98 (47%) mężczyźni. Szczegółowa struktura respondentów według płci i wieku została przedstawiona w tabeli 5.3. Ankieta zawierała łącznie 23 pytania, jednak z uwagi na fakt występowania pytań filtrujących, żadna z badanych osób nie odpowiadała na cały zestaw. Pytania występujące w ankiecie posiadają w większości przypadków formę zamkniętą, wielokrotnego wyboru lub oceny, gdzie każda osoba została zobligowana do określenia poziomu każdego z wymienionych czynników. Ponadto, badanie zostało uzupełnione o wypowiedź otwartą, w której respondenci mogli bardziej szczegółowo opisać swoje przemyślenia na temat działania prototypu rękawicy.

Tab. 5.3. Rozkład respondentów ankiety z podziałem na płeć i wiek (opracowanie własne)

			Płeć		Ogółem
			Kobieta	Mężczyzna	
Wiek	18 - 29	Liczebność	86	80	166
		%	77	82	79
	30 - 39	Liczebność	17	10	27
		%	15	10	13
	40 - 49	Liczebność	7	5	12
		%	6	5	6
	50 - 59	Liczebność	2	3	5
		%	2	3	2
	60 i więcej	Liczebność	0	0	0
		%	0	0	0
	Ogółem	Liczebność	112	98	210
		%	100	100	100

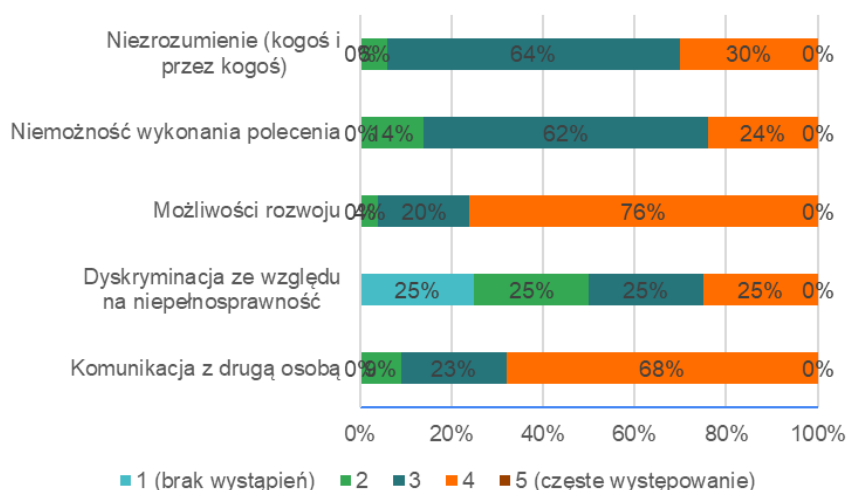
Ankieta została podzielona na kilka sekcji, spośród których każda z wyjątkiem pierwszej, była przeznaczona tylko dla określonej grupy osób. Kluczowymi osobami, z uwagi na cel niniejszej pracy, były osoby pracujące w przedsiębiorstwie produkcyjnym. Była to największa grupa, licząca 101 ankietowanych, z których 54 to osoby niesłyszące lub niedosłyszące. Tak liczna reprezentacja osób z niepełnosprawnościami słuchowymi stanowi silną podstawę do wyciągnięcia możliwie jak najbardziej prawidłowych, niespolaryzowanych wniosków.

Szczegółowa struktura badanych ze względu na branżę w jakiej wykonują swoje obowiązki zawodowe została przedstawiona w tabeli 5.4.

Tab. 5.4. Rozkład respondentów ankiety z podziałem na branże (opracowanie własne)

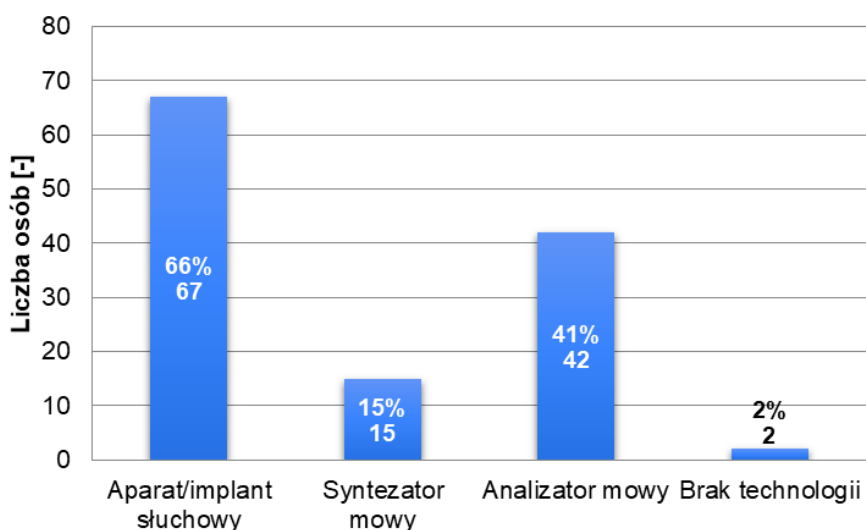
			Osoby słyszące	Osoby niesłyszące	Osoby niedosłyszające
Branża	Administracja	Liczebność	8	12	6
		%	9	19	16
	Bankowość	Liczebność	5	0	0
		%	6	0	0
	Usługi	Liczebność	9	2	6
		%	10	3	16
	IT	Liczebność	8	19	3
		%	9	30	8
	Edukacja	Liczebność	9	0	0
		%	10	0	0
	Służby mundurowe	Liczebność	3	0	0
		%	3	0	0
	Przedsiębiorstwo produkcyjne	Liczebność	47	31	23
		%	53	48	61
Ogółem		Liczebność	89	64	38
		%	100	100	100

Rysunek 5.10 przedstawia strukturę trudności z jakimi osoby niesłyszące i niedosłyszające zmagają się w życiu zawodowym. Czynnikiem, które wskazywały osoby z niepełnosprawnościami słuchowymi jako najczęściej występujące to możliwość rozwoju oraz komunikacja z drugą osobą. Z punktu widzenia tej grupy ankietowanych niezrozumienie (kogoś i przez kogoś) a także niemożność wykonania polecenia występuje w większości przypadków w umiarkowanym zakresie (ocena 3 w skali od 1 do 5). Dyskryminacja ze względu na niepełnosprawność została najbardziej równomiernie podzielona przez badanych. Osoby niesłyszące i niedosłyszające pracujące (obecnie lub w przeszłości) w przedsiębiorstwie produkcyjnym jako najczęściej występującą trudność wskazywały komunikację z drugą osobą (55% tej podgrupy ankietowanych oceniło tą trudność na wartość 4 w skali od 1 do 5).



Rys. 5.10. Rozkład napotykaných trudności w życiu zawodowym osób niesłyszących i niedosłyszących (opracowanie własne)

Na wykresie przedstawionym na rysunku 5.11 widać z jakich urządzeń i technologii korzystają najczęściej osoby z niepełnosprawnościami słuchowymi. 67 ankietowanych (66% tej grupy respondentów) korzysta lub korzystało z aparatu bądź implantu słuchowego a 42 badanych (41%) używało analizatora mowy. Najrzadziej występującymi urządzeniami były syntezatory mowy (15 osób). Dwie osoby nie wskazały żadnego urządzenia czy technologii, żadna osoba natomiast nie wskazała bardziej zaawansowanych technologicznie systemów, jak systemów wizyjnych czy inteligentnych form ubioru.



Rys. 5.11. Wykorzystywane przez osoby głuchonieme w życiu zawodowym technologie i urządzenia (opracowanie własne)

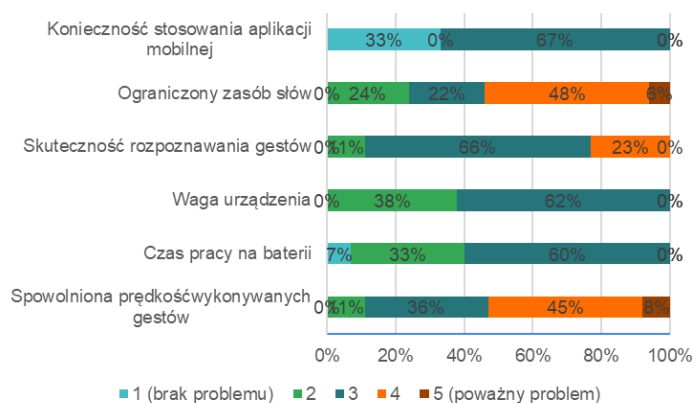
W pytaniu numer 10, ankietowani zostali zapytani czy potrafią posługiwać się językiem migowym (wyniki przedstawione na rysunku 5.12). Spośród osób niesłyszących i niedosłyszących 73% wskazało, że zna język migowy. W kolejnym pytaniu zapytano

o stopień znajomości najpopularniejszych języków migowych: Polskiego Języka Migowego (PJM), Brytyjskiego Języka Migowego (BSL), Amerykańskiego Języka Migowego (ASL) oraz międzynarodowego języka osób głuchych GESTUNO. Wszyscy respondenci wykazali co najmniej średnią znajomość Polskiego Języka Migowego (ocena 3 lub wyższa w skali od 1 do 5). Pozostałe języki zostały ocenione w różnym stopniu znajomości.



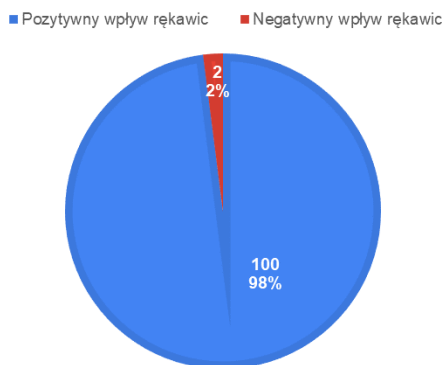
Rys. 5.12. Znajomość języka migowego wśród niesłyszących i niedosłyszących respondentów (opracowanie własne)

Kolejnym etapem było zaprezentowanie osobom niesłyszącym i niedosłyszącym działania prototypu rękawicy. Przedstawiona została specyfika fizyczna (hardware) oraz skrótowo objaśnione zostały aspekty programowe i komunikacyjne a także poruszanie się po aplikacji mobilnej. Następnie wykonane zostały gesty i przeprowadzona została ich klasyfikacja. Na rysunku 5.13 wykazane zostały najbardziej prawdopodobne potencjalne trudności związane z użytkowaniem rękawicy jakie zostały wskazane przez grupę docelową. Jako najbardziej problematyczne (ocenione co najmniej na 4 w skali od 1 do 5 przez ponad 50% tej grupy respondentów) zostały wykazane: ograniczony zasób słów oraz spowolniona prędkość wykonywania gestów. Innym potencjalnym problemem, który dla 23% ankietowanych stanowił poważniejszą trudność była skuteczność rozpoznawania gestów. Pozostałe czynniki, jakimi były: konieczność stosowania aplikacji mobilnej, waga urządzenia i czas pracy na baterii zostały przez większość ocenione jako neutralne (ocena 3 w skali od 1 do 5).



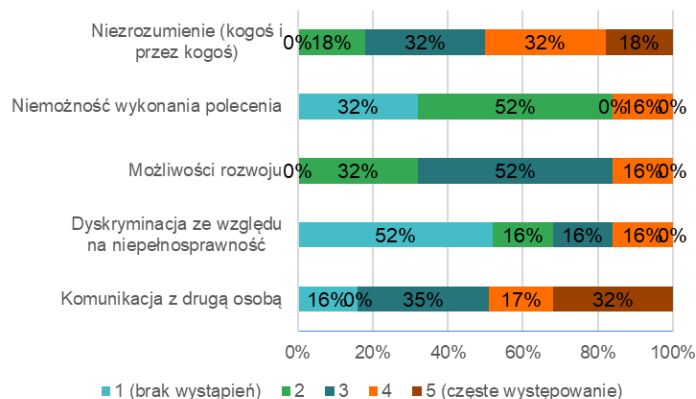
Rys. 5.13. Rozkład trudności związanych z użytkowaniem rękawicy przez osoby głuchonieme z podziałem na wagi (opracowanie własne)

Analizując odpowiedzi 54 respondentów należących do grupy osób niesłyszących lub niedosłyszących w środowisku produkcyjnym można zauważyć, że średnio o 10 punktów procentowych więcej z nich uważa za poważny problem spowolnioną prędkość rozpoznawanych słów oraz skuteczność ich rozpoznawania w stosunku do wszystkich ankietowanych posiadających niepełnosprawność słuchową. Ponadto 98% niesłyszących i niedosłyszących respondentów oceniło, że stosowanie rękawic interpretujących język migowy na tekst i mowę pozytywnie wpłynęłoby na ich życie zawodowe (rysunek 5.14).



Rys. 5.14. Wpływ inteligentnych rękawic na poprawę komunikacji w pracy osób z niepełnosprawnościami słuchowymi z punktu widzenia osób niesłyszących i niedosłyszących (opracowanie własne)

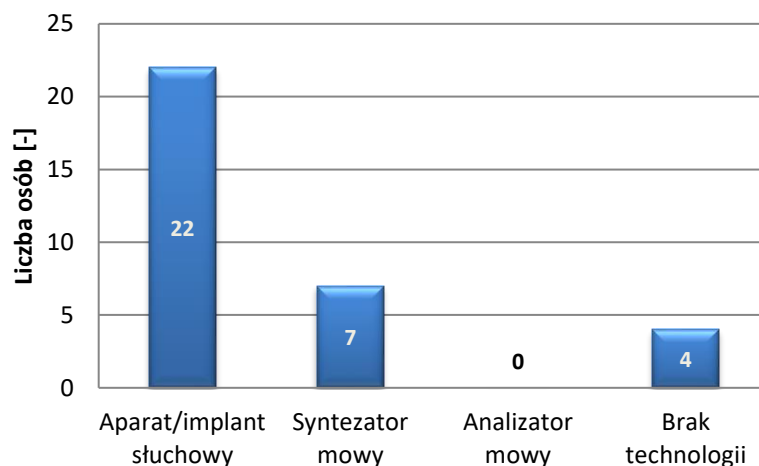
Na rysunku 5.15 przedstawiono trudności z jakimi zmagają się osoby niesłyszące i niedosłyszące przy wykonywaniu obowiązków zawodowych, z którymi współpracowały osoby słyszące. Najczęściej występującymi trudnościami wskazanymi przez respondentów były: nieporozumienie (kogoś lub przez kogoś) oraz komunikacja z drugą osobą. Możliwość rozwoju została wskazana przez 52% respondentów jako występująca w umiarkowanej częstotliwości (ocena 3 w skali od 1 do 5), natomiast niemożność wykonania polecenia i dyskryminacja ze względu na niepełnosprawność były określane jako najrzadziej występujące.



Rys. 5.15. Rozkład napotykaných trudności w życiu zawodowym osób niesłyszących i niedosłyszących z punktu widzenia osób słyszących (opracowanie własne)



Ta sama grupa osób określiła z jakich urządzeń i technologii korzystały osoby z uszkodzonym narządem słuchu. Jak widać na rysunku 5.16, najczęściej stosowanym urządzeniem był aparat lub implant słuchowy, wykorzystywany przez 22 osoby, następnie syntezytor mowy (7 osób). Analizator mowy nie był użytkowany przez żadną osobę, natomiast w 4 przypadkach nie była wykorzystywana żadna technologia ani urządzenie.



Rys. 5.16. Wykorzystywane przez osoby głuchonieme w życiu zawodowym technologie i urządzenia (opracowanie własne)

Rysunek 5.17 prezentuje strukturę wag potencjalnych trudności, związanych z użytkowaniem inteligentnej rękawicy, która została utworzona na podstawie odpowiedzi grupy respondentów złożonej z osób słyszących oraz niesłyszących lub niedosłyszących nieznających języka migowego. Do najpoważniejszych problemów zostały zaliczone: ograniczony zasób słów i skuteczność rozpoznawania gestów. Nieco mniejszą trudnością okazał się czas pracy na baterii oraz spowolniona prędkość wykonywanych gestów. W przypadku konieczności stosowania aplikacji mobilnej, 66% respondentów uznało, że nie stanowi to żadnego problemu, natomiast 11% uznało, iż może to stanowić problem.



Rys. 5.17. Rozkład trudności związanych z użytkowaniem rękawicy przez osoby głuchonieme z podziałem na wagi z punktu widzenia osób słyszących (opracowanie własne)

Spośród 47 słyszących ankietowanych pracujących w przedsiębiorstwie produkcyjnym 56% uznało, że czas pracy na baterii jest stosunkowo poważnym problemem (ocena co najmniej 4 w skali od 1 do 5). Drugą najczęściej wskazywaną trudnością pod kątem wagi problemu była skuteczność rozpoznawania gestów (48% respondentów) oraz ograniczony zasób słów (46% ankietowanych). Z tej grupy osób, 88% oceniło przydatność rękawicy w wykonywaniu obowiązków zawodowych przez osoby głuchonieme, natomiast rozpatrując całą grupę osób słyszących i nieznających języka migowego 12% oceniło wykorzystywanie inteligentnych rękawic jako nieprzynoszące ułatwienia niepełnosprawnym (rysunek 5.18).



Rys. 5.18. Wpływ inteligentnych rękawic na poprawę komunikacji w pracy osób z niepełnosprawnościami słuchowymi z punktu widzenia osób słyszących (opracowanie własne)

## 5.4. Dyskusja i wnioski

W odniesieniu do osiągniętych wyników zaprezentowanych w podrozdziale 5.2, na pierwszy plan wysuwa się niestabilność związana z wytrzymałością połączeń elektrycznych. Ciągłe ruchy w przestrzeni, niejednokrotnie gwałtowne a także kontakt z ciałami obcymi wpisują się w specyfikę korzystania z rękawicy w środowisku produkcyjnym. W takim przypadku połączenia przewodów z mikrokontrolerem i sensorami powinny zostać z pewnością lepiej zabezpieczone przed czynnikami zewnętrznymi, aby nie doprowadzać do uszkodzeń i zachować stabilność odczytów.

Kolejne wersje prototypu rękawicy powinny zostać wyposażone w baterię litowo-polimerową o pojemności od 300mAh do 500mAh. Zastosowanie niewielkich gabarytowo źródeł energii znacząco poprawi komfort użytkownika, a zastosowanie ogniwo o jednej z wyżej wymienionych pojemności, przy wykorzystanych energooszczędnych podzespołach, pozwoli na nieprzerwaną pracę do około 5 godzin. Aby zachować długą żywotność akumulatora należałoby zastosować ładowarkę o stosunkowo niskiej mocy. Ważną zmianą, w kwestii samej rękawicy, będzie wykorzystanie dzianin syntetycznych, stosowanych, np. w bieliznie i odzieży

sportowej. Na rynku można spotkać się z wieloma typami materiałów (poliamid, elastan i polipropylen), dlatego konieczne będzie przeanalizowanie wszystkich właściwości wpływających na komfort użytkowania, m.in. elastyczność, stopień przepuszczania powietrza, wytrzymałość na rozciąganie i uderzenia mechaniczne.

Ważnym wyzwaniem w następnych etapach pracy nad urządzeniem będzie stworzenie drugiej, symetrycznej rękawicy oraz zaimplementowanie w niej tego samego oprogramowania. Dużo większym wyzwaniem w procesie dodania drugiego urządzenia będzie natomiast zmiana logiki dla programu aplikacji mobilnej, która będzie musiała obsługiwać dwa moduły Bluetooth oraz analizować dane z dwóch nadajników. Integracja dwóch bliźniaczych urządzeń z aplikacją mobilną jest dużym wyzwaniem, jednak daje dużo większe możliwości. Otrzymane większe paczki danych (wartości charakterystycznych oraz ugięcia palców), pozwolą dużo lepiej rozróżniać poszczególne gesty, ze względu na większe możliwości kombinacji liczb w macierzy parametrów.

Średnie czasy związane z wykonywaniem obliczeń na mikrokontrolerze są stabilne, dzięki czemu wszelkie niedokładności i braki danych mogą zostać w stosunkowo prosty sposób zaproksymowane przez algorytmy, jako że czasy te stanowią maksymalnie 12% całkowitego czasu trwania ramki gestu. Większą trudność stanowi natomiast szybkość wykonywania klasyfikacji przez klasyfikator. O ile bowiem przy 50 słowach w słowniku, czasy te osiągnięte przez aplikację mobilną są pomijalnie małe, o tyle już przy 4000 słowach pojawiłoby się zauważalne opóźnienie bądź utrata części danych otrzymywanych z mikrokontrolera rękawicy. W związku z tym efektem, przy posiadaniu tak obszernej bazy słów należałoby wykonać optymalizację klasyfikatora lub definicji atrybutów opisowych.

Przyglądając się współczynnikowi skuteczności całościowej można powiedzieć, że wynik na poziomie 92% jest wysoki, aczkolwiek z pewnością proces wymaga jeszcze optymalizacji. W celu zwiększenia skuteczności średniej dla wszystkich gestów ze słownika można zmodyfikować (tu: zwiększyć) wartości błędów granicznych, przez co zmaleje liczba niezakwalifikowanych do danej klasy słów. Trzeba jednak pamiętać, że jest to rozwiązanie, które może nieść za sobą zmniejszenie skuteczności jednostkowej oraz pogorszenie się parametrów czułości lub specyficzności dla konkretnych wyrażeń, ponieważ zmniejszą się granice między klasami. W tym elemencie trzeba zachować szczególną ostrożność i po każdorazowych zmianach wykonać co najmniej testy offline na bazie zgromadzonych wcześniej danych. Ponownie może się okazać, że jedynym rozwiązaniem dla zwiększenia skuteczności całościowej nie pogarszając przy tym pozostałych parametrów klasyfikacji będzie optymalizacja lub zmiana klasyfikatora.

Alternatywnym rozwiązaniem, poprawiającym skuteczność rozpoznawania gestów jest zwiększenie liczby obliczanych wartości charakterystycznych. Duża liczba różnych parametrów potęguje możliwości kombinacyjne, przez co każdy z gestów może być bardziej unikatowy (nie będzie mylony z innym). Kluczowym może być skorzystanie z odczytów innych sensorów, które mogłyby zostać zintegrowane w układzie sterującym na rękawicy oraz wykorzystanie tych danych do obliczeń wartości charakterystycznych. Jednak takie rozwiązanie niesie za sobą kilka następstw. Większa liczba parametrów oznacza wolniejszy transfer danych przez Bluetooth, a także wymaga stworzenia macierzy o większych wymiarach danych w bazie SQL. Przeszukiwanie dużych paczek danych w bazach, może wprowadzić opóźnienia, które miałyby negatywny wpływ na pracę całego urządzenia.

Innym, ciekawym sposobem poprawy efektywności odczytywania słów w języku migowym jest optymalizacja algorytmu przeszukującego bazę SQL. Można wprowadzić elementy sortowania po parametrach charakterystycznych, w taki sposób, aby obliczać błędy tylko dla słów z danymi najbardziej zbliżonymi do danych odczytywanych w czasie rzeczywistym. Ważne, aby przy każdych zmianach algorytmu sprawdzać, czy nowe rozwiązanie daje większe korzyści. W takim przypadku trzeba przeprowadzać testy szybkości przeszukiwania całej bazy danych.

Tworzenie algorytmów klasyfikacyjnych oraz ekstrakcja atrybutów opisowych dla tak złożonego procesu jakim jest komunikacja w języku migowym należy do trudnych zadań. Osiągnięcie całkowitej płynności i prędkości wykonywania gestów przy zachowaniu bardzo wysokiej skuteczności rozpoznawalności w czasie rzeczywistym wiąże się z koniecznością wykonania bardzo wielu testów i optymalizacji. Otrzymane w niniejszej pracy wyniki stanowią dobrą podstawę do głębszej analizy tematu, jednak z pewnością nie pozwalają na implementację rozwiązania w takiej formie w warunkach docelowych, a więc w przemyśle w warunkach produkcyjnych.

W urządzeniu komunikacji głosowej jednym z najbardziej czasochłonnych etapów jest dodawanie nowych wyrażen do słownika. Proces ten opiera się na wymagających próbach zręcznościowych oraz zaawansowanych obliczeniach matematycznych. Innowacyjnym rozwiązaniem byłoby zastosowanie algorytmu opartego na sieciach neuronowych. Takie rozwiązanie pozwoliłoby wprowadzić element sztucznej inteligencji do rękawicy, co przełożyłoby się na nowe możliwości, np. uczenie się nowych słów oraz doskonalenie ich atrybutów z każdym kolejnym gestem. Z czasem sieć użytkowników mogłaby w szybki sposób uzupełnić słownik aplikacji o brakujące wyrażenia.

Na podstawie zebranych wyników w ankiecie widać, że wśród osób niesłyszących i niedosłyszących istnieje poważny problem przede wszystkim związany z komunikacją z drugą osobą a także z możliwościami rozwoju zawodowego. Te same trudności okazały się kluczowe dla osób z niepełnosprawnościami słuchowymi pracującymi w przedsiębiorstwie produkcyjnym. Istnieje zatem duży potencjał w stworzeniu i skutecznym wdrożeniu prototypu, który usprawniłby takim osobom wykonywanie obowiązków zawodowych.

Warto także zwrócić uwagę na problemy, które są obecne w aktualnej wersji prototypu urządzenia. Dla osób niesłyszących i niedosłyszących są to przede wszystkim ograniczony zasób słów oraz spowolniona prędkość wykonywanych gestów. Widać zatem u tych osób szczególnie silne przywiązanie do swobody wypowiedzania się, którą posiadają w swoim środowisku i nie chcieliby z niej rezygnować. Zatem dla grupy docelowej użytkowników, takie rozwiązanie nie mogłoby wprowadzać dodatkowych ograniczeń. Dodatkowym aspektem, na który należy zwrócić uwagę jest także brak możliwości konstruowania pełnych zdań w obecnej wersji prototypu rękawicy, co stanowi istotne ograniczenie. W finalnej wersji urządzenia z pewnością należałoby umożliwić użytkownikom swobodne formułowanie wypowiedzi i interpretowanie gramatyki pozycyjnej języka migowego na gramatykę języka narodowego.

Warto porównać także największe trudności jakie w obecnym rozwiązaniu widzą osoby słyszące i nieznające języka migowego. Są to przede wszystkim, podobnie jak dla głuchoniemych, ograniczony zasób słów a także skuteczność rozpoznawania gestów, która była wcześniej wskazywana przez tą grupę respondentów. Takie trudności natomiast jak waga urządzenia, czas pracy na baterii czy konieczność stosowania aplikacji mobilnej, które dla osób głuchoniemych stanowiły mały lub neutralny stopień trudności, dla osób słyszących niejednokrotnie zostały wskazane jako problem poważny. Widać zatem, że aspekty czysto sprzętowe są inaczej odbierane przez te dwie grupy ankietowanych, czego przyczyną może być zarówno różnica w świadomości i doświadczeniu w użytkowaniu podobnej klasy sprzętów jak i stawianie elementów ukierunkowanych na komunikację z drugim człowiekiem na różnym poziomie istotności.

## 6. Podsumowanie

Sytuacja osób niesłyszących na rynku pracy nie należy do najłatwiejszych, czego dowodem jest wysoki odsetek tej grupy osób, która nawet nie próbuje podjąć się pracy zarobkowej. Przyczyn takiego stanu jest wiele, a najczęściej pojawiającym się w literaturze jest problem w komunikacji z osobami słyszącymi, nieznanymi języka migowego. Oczywistym jest fakt, że nie da się poznać osobowości czy charakteru człowieka bez rozmowy, dlatego tak kluczowe jest odnalezienie metody rozwiązania problemu osób głuchych. Dotychczasowe techniki wspomagania niesłyszących opierały się na mało zaawansowanych środkach, jak np. płatne wideokonferencje z tłumaczem języka migowego. Co prawda w ciągu ostatnich kilkudziesięciu lat powstało kilka zaawansowanych technicznie prototypów urządzeń, które podjęły próbę odczytywania miganych gestów i tłumaczenia ich na języki narodowe, jednak nie udało się dopracować tych wynalazków do takiego stopnia, aby mogły być wprowadzone na rynek.

Głównym celem realizacji niniejszego projektu było stworzenie urządzenia komunikacji głosowej, wspomagającej osoby niesłyszące oraz słyszące we wzajemnej, bezpośredniej wymianie informacji. Istotnym wymaganiem była odporność na wszelkie zagrożenia, wynikające z zastosowania urządzenia w środowiskach produkcyjnych z zachowaniem jego pełnej funkcjonalności. Zakłócenia w postaci hałasu, silnych fal magnetycznych, radiowych, oraz niewielkie obszary stanowisk pracy spowodowały, że zdecydowano się skonstruować kompaktowe, lekkie oraz mobilne urządzenie w postaci rękawicy, zabudowanej mikroelektroniką.

Wynikiem pracy zespołu autorów jest prototyp rękawicy wraz z aplikacją dla urządzeń mobilnych. Kluczowymi punktami procesu budowy urządzenia był dobór odpowiednich elementów elektroniki oraz stworzenie algorytmu funkcjonowania oprogramowania rękawicy i aplikacji (m.in. algorytmy klasyfikacji, wyszukiwania). Przemyślany wybór układów sterujących dla rękawicy pozwolił osiągnąć zamierzone cele wydajnościowe, a innowacyjne rozwiązania programowe oraz liczne analizy pozwoliły doprowadzić urządzenie komunikacji głosowej do przeprowadzania rozpoznawalności gestów na wysokim poziomie.

Napisana na potrzeby projektu aplikacja mobilna jest nieodłącznym elementem całego urządzenia. Wykorzystanie bazy danych SQL pozwala efektywnie przeszukiwać kolejne rekordy danych. Do tej pory w słowniku aplikacji zapisano 50 słów, z których większość dotyczy środowiska produkcyjnego. Docelowym rozwiązaniem ma być algorytm uczący się nowych gestów. Wprowadzenie takiej funkcjonalności pozwoliłoby tworzyć słownik gestów w każdym języku migowym świata przy pomocy aktywnych użytkowników.

Istotnym problemem tłumaczeń z języka migowego jest obecność odmiennej gramatyki, co powoduje tworzenie zdań w szyku przestawnym przez osobę głuchoniemą. Aby móc poprawnie gramatycznie przekazywać informacje od docelowych użytkowników opisywanego urządzenia, należałoby uwzględnić także ten element w algorytmach rozpoznawania gestów.

Przy relatywnie niedużych niedokładnościach urządzenia udało się osiągnąć także inny kluczowy element, który występuje w prawie każdym dostępnym kanale komunikacyjnym bezpośredniego kontaktu, a więc pracę w czasie rzeczywistym. Aby nie powodować dodatkowego dyskomfortu i presji na użytkownika, wymiana komunikatów powinna przebiegać bez zauważalnych opóźnień i taką właściwość posiada skonstruowane urządzenie, pomimo wykonywania wielu złożonych obliczeń, bezprzewodowej wymiany danych oraz przeprowadzania procesu klasyfikacji.

W początkowej fazie projektu autorzy byli przekonani o tym, że głównym celem zastosowania urządzenia będzie wspomaganie osób głuchych w codziennym życiu oraz pracy. Przy współpracy z osobami niesłyszącymi ten punkt widzenia znacząco się zmienił. Okazało się, że konstruowane urządzenie komunikacji głosowej będzie raczej wspomagało osoby słyszące w zrozumieniu niesłyszących współpracowników.

Aktywizacja osób z niepełnosprawnościami słuchowymi do podjęcia pracy zawodowej jest istotnym społecznie wyzwaniem, którego poziom istotności będzie z biegiem czasu narastał z uwagi na zwiększanie się liczby osób niesłyszących i niedosłyszących. Efekty podjętej próby stworzenia urządzenia komunikacji głosowej dla osób głuchoniemych w warunkach produkcyjnych nie mogą zostać uznane za ostateczny wzór systemu, natomiast stanowią one solidną podstawę, na której można oprzeć konstrukcję algorytmów i strukturę sprzętową docelowego urządzenia. Świadczyć o tym mogą opracowane, innowacyjne algorytmy ekstrakcji atrybutów opisowych i ich wykorzystanie w późniejszym procesie klasyfikacji, których implementacja w stworzonej aplikacji mobilnej pozwoliła na swobodne wykonywanie gestów w języku migowym i tłumaczenie ich na język narodowy.

## Literatura

AGARWAL T., (2017), *Wi-Fi Working Principle, Types of Technologies and Application*, <https://www.elprocus.com/how-does-wifi-work/> (dostęp: 27.09.2019)

ANDRZEJEWSKI W., (2008), *Klasyfikacja*, [https://www.cs.put.poznan.pl/wandrzejewski/wp-content/uploads/aed\\_wsnhid/Klasyfikacja.pdf](https://www.cs.put.poznan.pl/wandrzejewski/wp-content/uploads/aed_wsnhid/Klasyfikacja.pdf) (dostęp: 10.09.2019)

BEDNAREK A., (2014), *Xbox One: jak działa Kinect 2.0*, <https://www.komputerswiat.pl/poradniki/gaming/xbox-one-jak-dziala-kinect-20/e1414vn> (dostęp: 13.10.2019)

BIELAK T., BORKOWSKI I., DAWIDZAK-KŁADOCZNA M., GIELATA I., GOROŃSKA A., GÓRAL-PÓŁROLA J., JELC A., PACUŁA J., PASZEK J., ROMANISZYN-ZIOMEK L., WĄDOLNY-TATAR K., ZAKRZEWSKA S., (2014), *Poetyka gestów. Język – Literatura – Kultura*, Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej, Bielsko Biała

BOTLAND, (2019a), *Moduł Bluetooth HC-05 GW-040*, <https://botland.com.pl/pl/produkty-wycofane/5244-modul-bluetooth-hc-05-gw-040.html> (dostęp: 15.10.2019)

BOTLAND, (2019b), *Moduł WiFi ESP8266 Black – 3 GPIO, 1MB, PCB antena*, <https://botland.com.pl/pl/moduly-wifi/4527-modul-wifi-esp-01-esp8266-black-3-gpio-1mb-pcb-antena.html> (dostęp: 15.10.2019)

BOTLAND, (2019c), *Moduł Bluetooth 4.0 BLE - HM - 10 - 3,3/5V*, <https://botland.com.pl/pl/moduly-bluetooth/8591-modul-bluetooth-40-ble-hm-10-33v5v.html> (dostęp: 15.10.2019)

BOTLAND, (2019d), *Leap Motion Controller - precyzyjny czujnik gestów*, <https://botland.com.pl/pl/czujniki-gestow/13755-leap-motion-controller-precyzyjny-czujnik-gestow-857192004084.html> (dostęp: 10.10.2019)

BOTLAND, (2019e), *Czujnik ugięcia 73x6,3mm – SparkFun*, <https://botland.com.pl/pl/czujniki-nacisku/1640-czujnik-ugiecia-73x63mm-sparkfun.html> (dostęp: 13.10.2019)

BRIGANCE J., (2017), *WiFi vs Bluetooth vs BLE, choosing the right IoT tech*, <https://appdeveloperomagazine.com/wifi-vs-bluetooth-vs-ble,-choosing-the-right-iot-tech/> (dostęp: 07.10.2019)

BUENAFLORES D., (2019), *What's Difference Between Bluetooth Versions 2.x, 3.x, 4.x and 5.x*, <https://thedroidguy.com/2019/09/whats-difference-between-bluetooth-versions> (dostęp: 27.09.2019)

BUTKIEWICZ A., BOGUSZYŃSKA I., BRZEGOWA-JABCZYK A., GRZEŚKOWIAK A., KEMPCZYŃSKA W., PIETRUS-RAJMAN A., SIURDYBAN M., (2014b), *Sytuacja i możliwości aktywizacji głuchych*, Dobre Kadry, Wrocław

BUTKIEWICZ A., CZAJKOWSKA-KISIL M., DUNAJ M., KOWALSKI P., LASKOWSKA - KLIMCZEWSKA A., ŁACHETA J., MACIEJESKA - ROCZAN K., MARGANIEC B., SACHA A., SAK M., SIEPKOWSKA A., SZCZYGIELSKA M., (2014a), *Sytuacja osób głuchych w Polsce*, Biuro Rzecznika Praw Obywatelskich, Warszawa



CHMIELNICKI W., (2012), *Efektywne metody selekcji cech i rozwiązywania problemu wieloklasowego w nadzorowanej klasyfikacji danych*, Kraków

COXWORTH B., *Student-designed aid for the deaf converts speech to AR captions*, <https://newatlas.com/live-time-closed-captioning-system/40078/> (dostęp 13.11.2019)

DFROBOT, (2019), *DFRobot CurieNano – A mini Genuino Arduino101 Board SKU DFR0453*, [https://wiki.dfrobot.com/DFRobot\\_CurieNano\\_\\_A\\_mini\\_Genuino\\_Arduino\\_101\\_Board\\_SKU\\_DFR0453](https://wiki.dfrobot.com/DFRobot_CurieNano__A_mini_Genuino_Arduino_101_Board_SKU_DFR0453) (dostęp 15.10.2019)

DIKS K., (2007), *Klasyfikacja. Sformułowanie problemu, metody klasyfikacji, kryteria oceny metod klasyfikacji*

DUNAJ M. (2015), *Polski język migowy*, <https://rownosc.info/dictionary/polski-jezyk-migowy> (dostęp: 26.09.2019)

GRABOWSKA-ORZAŁA M., JÓZEFIAK J., LEMIROWSKA-WRONKA J., PODGÓRNI K., (2012), *Pl Zapisać ciszę. Podręcznik do nauki języka migowego z wykorzystaniem zapisu Sign Writing*, Nowy Sącz

GAŁKOWSKI T., KUNICKA-KAISERL., SMOLEŃSKA J., (1978), *Psychologia dziecka głuchego*, PWN, Warszawa

GRZYB M., (2017), *Wybór odpowiedniego algorytmu. Część 2 – algorytmy klasyfikacyjne*, [<https://mateuszgrzyb.pl/wybor-odpowiedniego-algorytmu-czesc-2-algorytmy-klasyfikacyjne/>] (dostęp: 23.09.2019)

GUS, (2012), *Osoby niepełnosprawne na rynku pracy w 2011r. Informacje i opracowania statystyczne*, Zakład Wydawnictw Statystycznych, Warszawa

HANDSPEAK, (2019), *Identify different classes of classifiers*, <https://www.handspeak.com/learn/index.php?id=20> (dostęp: 13.10.2019)

HARASYMCZUK M., (2017), *Python 3 from None to Machine Learning*, <http://python.astrotech.io/machine-learning/bayes.html> (dostęp: 27.09.2019)

KOZICKA M., (2011), *Klasyfikacja uszkodzeń słuchu*, <http://www.sds24.pl/pojecie-gluchoty> (dostęp: 26.09.2019)

MAN M., (2002), *Bluetooth and Wi-Fi. Understanding these two technologies and how they can benefit you*, Socket Communications, Newark

MEDGADGET, (2014), *MotionSavvy UNI Sign Language Interpreter Gives a Voice to the Deaf and Hard of Hearing*, <https://www.medgadget.com/2014/10/motionsavvy-uni-sign-language-interpreter-gives-a-voice-to-the-deaf-and-hard-of-hearing.html> (dostęp: 10.10.2019)

MERIAUDEAU F., (2012), *Sign Language Translator using Microsoft Kinect*, <https://www.semanticscholar.org/paper/Sign-Language-Translator-using-Microsoft-Kinect-M%C3%A9riaudeau/165a1a7c529f51b91ae587496d41603e560d1fe9/> (dostęp: 14.10.2019)

MICHALAK Ł., (2013), *Polak potrafi, czyli KinectTranslator. Kinect jako tłumacz języka migowego*, <https://gadzetomania.pl/3103,polak-potrafi-czyli-kinecttranslator-kinect-jako-tlumacz-jezyka-migowego> (dostęp: 11.10.2019)

- MIESZKOWICZ M., (2012), *Język migowy a migany w komunikacji osób niesłyszących*, <http://poradnik-logopedyczny.pl/komunikacja-alternatywna/komunikacja-alternatywna/247/jezyk-migowy.html> (dostęp: 22.09.2019)
- MIRAGE, (2018), *Rękawice bawełniane RWKB*, <http://mirage-bhp.pl/products/rekawice-bawelniane-rwkb/> (dostęp: 15.10.2019)
- PŁUSKA K., (2017), *Kanał, szumy i bariery... Komunikacja werbalna*, <https://www.katarzynapluska.pl/kanal-szumy-i-bariery-komunikacja-werbalna/> (dostęp: 18.09.2019)
- SAK M., (2012), *4 kroki. Wsparcie osób niesłyszących na rynku pracy II*, Państwowy Fundusz Rehabilitacji Osób Niepełnosprawnych, Warszawa
- SCIENCEDAILY, (2014), *Smartfone understands hand gestures*, <https://www.sciencedaily.com/releases/2014/10/141008131613.htm> (dostęp: 10.10.2019)
- SĘKOWSKA Z., (1985), *Pedagogika specjalna*, Państwowe Wydawnictwo Naukowe, Warszawa
- STRAUSS K., (2014), *Tech Tackles Sign Language – MotionSavvy*, <https://www.forbes.com/sites/karstenstrauss/2014/10/27/tech-tackles-sign-language-motionsavvy/#b04608678627> (dostęp: 11.10.2019)
- PRERNA H., (2017), *SignAloud Gloves Convert Sign Language into Spoken Words*, <https://www.thetechy.com/signaloud-gloves-convert-sign-language/> (dostęp 05.10.2019)
- THOMPSON A., (2017), *This Hacked Together Glove Can Translate Sign Language*, <https://www.popularmechanics.com/technology/gear/a27292/glove-translate-sign-language/> (dostęp: 11.09.2019)
- THORN T., (2014), *WiFi Direct: what it is and why you should care*, <https://www.techradar.com/news/phone-and-communications/mobile-phones/wi-fi-direct-what-it-is-and-why-you-should-care-1065449> (dostęp: 28.09.2019)
- WANSHEL E., (2016), *Students Invented Gloves That Can Translate Sign Language Into Speech And Text*, [https://www.huffpost.com/entry/navid-azodi-and-thomas-prior-signaloud-gloves-translate-american-sign-language-into-speech-text\\_n\\_571fb38ae4b0f309baeee06d?](https://www.huffpost.com/entry/navid-azodi-and-thomas-prior-signaloud-gloves-translate-american-sign-language-into-speech-text_n_571fb38ae4b0f309baeee06d?) (dostęp: 10.09.2019)

# Spis Rysunków

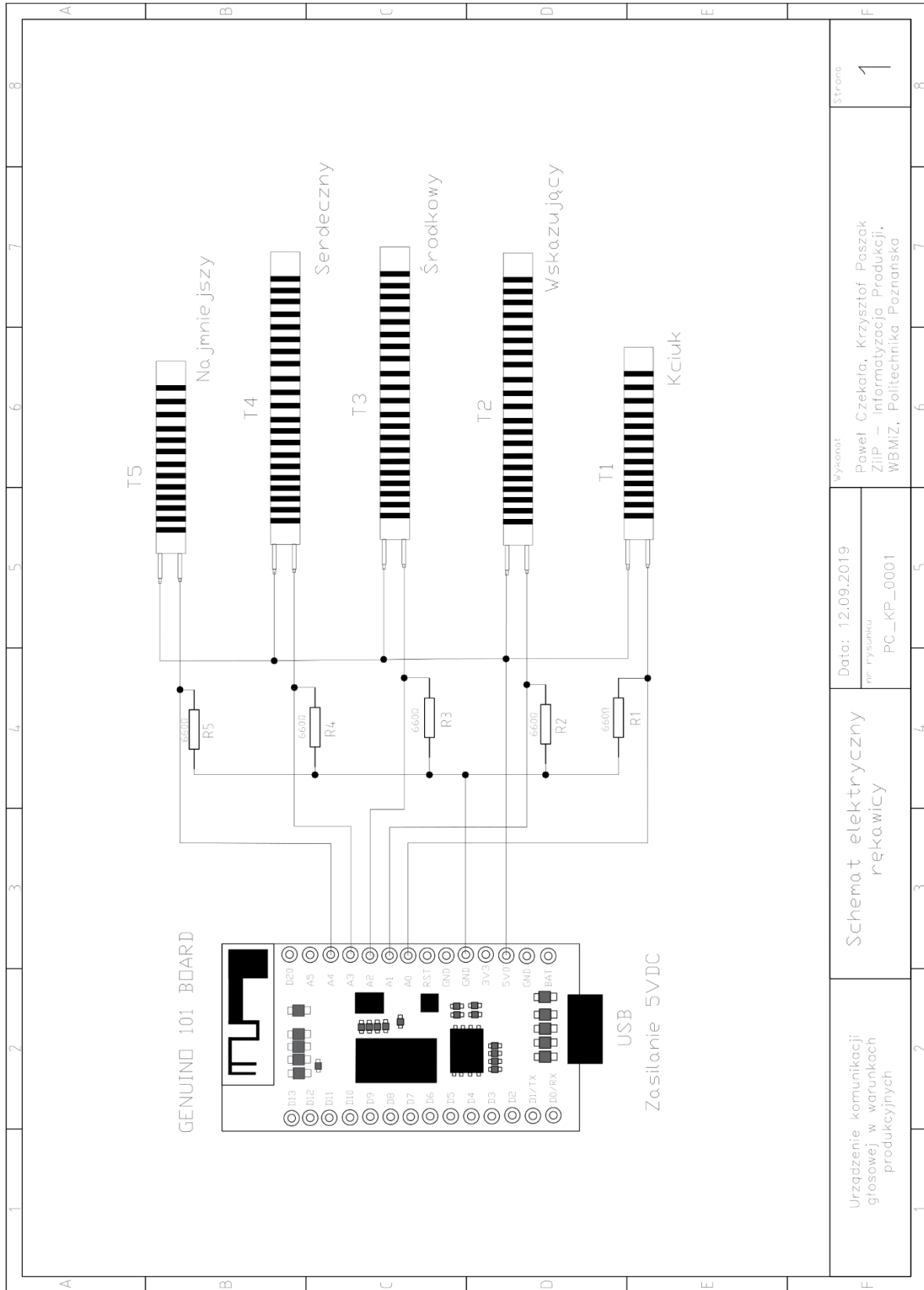
Rys. 1.1. Przepływ komunikatu między nadawcą a odbiorcą (opracowanie własne na podstawie (Płuska, 2017))	7
Rys. 3.1. Rozkład osób z niepełnosprawnością słuchową na rynku pracy (opracowanie własne na podstawie (GUS, 2012))	9
Rys. 3.2. Proces tworzenia klasyfikatora (opracowanie własne)	23
Rys. 3.3. Schemat procesu klasyfikacji (opracowanie własne)	24
Rys. 3.4. Obrazowy przykład działania drzewa decyzyjnego (opracowanie własne)	25
Rys. 3.5. Przykład działania klasyfikatora Bayesa (opracowanie własne na podstawie (Harasymczuk, 2019))	26
Rys. 3.6. Przykład działania klasyfikatora najbliższego sąsiedztwa (opracowanie własne na podstawie (Harasymczuk, 2019))	27
Rys. 3.7. Porównanie strategii OVR i OVO (opracowanie własne na podstawie (Chmielnicki, 2012))	28
Rys. 3.8. Wizualizacja koncepcji KinectTranslator (Mériaudeau, 2012)	30
Rys. 3.9. MotionSavvy's UNI, urządzenie do rozpoznawania gestów w języku migowym (Medgadget, 2014)	31
Rys. 3.10. SignAloud, prototyp rękawicy studentów z Uniwersytetu w Waszyngtonie (Perna, 2017)	32
Rys. 3.11. Prototyp rękawicy naukowców z San Diego (Thompson, 2017)	32
Rys. 4.1. Bawełniana rękawica (Mirage, 2018)	38
Rys. 4.2. Intel Curie Quark ARC (Dfrobot, 2019)	39
Rys. 4.3. Czujniki ugięcia firmy SparkFun (Botland, 2019e)	40
Rys. 4.4. Ogólny schemat kooperacji mikrokontrolera na rękawicy z aplikacją mobilną (opracowanie własne)	42
Rys. 4.5. Struktura danych w ramce komunikacyjnej (opracowanie własne)	43
Rys. 4.6. Struktura procesu pobierania i obróbki danych przez aplikację mobilną (opracowanie własne)	44
Rys. 4.7. Ogólna struktura rozkładu atrybutów opisowych (opracowanie własne)	45
Rys. 4.8. Szczegółowa struktura atrybutów opisowych w ramach komunikacyjnych (opracowanie własne)	45
Rys. 4.9. Schemat blokowy dla funkcji przeszukującej bazę słów (opracowanie własne)	47
Rys. 5.1. Prototyp rękawicy urządzenia komunikacji głosowej (opracowanie własne)	51
Rys. 5.2 Aplikacja mobilna a. Ekran główny b. Ekran rozpoznawania mowy (opracowanie własne)	52
Rys.5.3 Aplikacja mobilna a. Ekran menu b. Baza słów w aplikacji (opracowanie własne)	53
Rys.5.4 Aplikacja mobilna a. Słownik administratora b. Ekran ustawień (opracowanie własne)	54
Rys. 5.5 Aplikacja mobilna a. Ekran „O aplikacji” b. Ekran o autorach (opracowanie własne)	54
Rys. 5.6. Uszkodzenie połączenia elektrycznego rękawicy (opracowanie własne)	57
Rys. 5.7. Rozkład czasów obliczeń atrybutów opisowych i klasyfikacji gestu dla rosnącej liczby gestów (opracowanie własne)	57
Rys. 5.8. Rozkład błędów dla dostępnych gestów w słowniku (opracowanie własne)	57
Rys. 5.9. Udział błędnych odczytów gestów w całej próbie (opracowanie własne)	59
Rys. 5.10. Rozkład napotykanymi trudności w życiu zawodowym osób niesłyszących i niedosłyszących (opracowanie własne)	62
Rys. 5.11. Wykorzystywane przez osoby głuchonieme w życiu zawodowym technologie i urządzenia (opracowanie własne)	62

Rys. 5.12. Znajomość języka migowego wśród niesłyszących i niedosłyszących respondentów (opracowanie własne).....	63
Rys. 5.13. Rozkład trudności związanych z użytkowaniem rękawicy przez osoby głuchonieme z podziałem na wagi (opracowanie własne).....	63
Rys. 5.14. Wpływ inteligentnych rękawic na poprawę komunikacji w pracy osób z niepełnosprawnościami słuchowymi (opracowanie własne).....	64
Rys. 5.15. Rozkład napotykaných trudności w życiu zawodowym osób niesłyszących i niedosłyszących z punktu widzenia osób słyszących (opracowanie własne) .....	64
Rys. 5.16. Wykorzystywane przez osoby głuchonieme w życiu zawodowym technologie i urządzenia (opracowanie własne) .....	65
Rys. 5.17. Rozkład trudności związanych z użytkowaniem rękawicy przez osoby głuchonieme z podziałem na wagi z punktu widzenia osób słyszących (opracowanie własne) .....	65
Rys. 5.18. Wpływ inteligentnych rękawic na poprawę komunikacji w pracy osób z niepełnosprawnościami słuchowymi z punktu widzenia osób słyszących (opracowanie własne) .....	66

## Spis tabel

Tab. 3.1. Zestawienie parametrów popularnych modułów komunikacji bezprzewodowej (opracowanie własne na podstawie (Botland, 2019a, 2019b, 2019c)).....	20
Tab. 3.2. Porównanie popularnych urządzeń rozpoznających gesty (opracowanie własne) ...	20
Tab. 4.1. Plan prac wykonania prototypu rękawicy dla osób głuchoniemych (opracowanie własne) .....	36
Tab. 4.2. Zestawienie słów i wyrażeń słownika aplikacji mobilnej (opracowanie własne) ....	46
Tab. 4.3. Wizualizacja danych w tabeli bazy SQL (opracowanie własne) .....	46
Tab. 4.4. Schematyczny zapis rozkładu próbek przy wyznaczaniu współczynników czułości i specyficzności (opracowanie własne na podstawie (Chmielnicki, 2012, str. 38.39)).....	48
Tab. 5.1. Rozkład próbek dla klasy „stan” .....	58
Tab. 5.2. Rozkład wszystkich odczytów gestów (opracowanie własne) .....	59
Tab. 5.3. Rozkład respondentów ankiety z podziałem na płeć i wiek (opracowanie własne) .	60
Tab. 5.4. Rozkład respondentów ankiety z podziałem na branże (opracowanie własne) .....	61

# Załącznik 1. Schemat elektryczny urządzenia



Urządzenie komunikacji głosowej w warunkach produkcyjnych	Schemat elektryczny ręka.wicy	Data: 12.09.2019 nr rysunku PC_KP_0001	Wykonat Pawel Czekała, Krzysztof Paszak ZiIP – Informatyzacja Produkcji, WBMiZ, Politechnika Poznańska	Strona 1
---	----------------------------------	--	---	-------------

## Załącznik 2. Ankieta

### Sekcja: Podstawowe informacje

1. Płeć
  - a) Kobieta
  - b) Mężczyzna
  
2. Wiek
  - a) Poniżej 18
  - b) 18 – 29
  - c) 30 – 39
  - d) 40 – 49
  - e) 50 – 59
  - f) 60 i więcej
  
3. Wykształcenie
  - a) Podstawowe
  - b) Gimnazjalne
  - c) Zasadnicze zawodowe
  - d) Średnie
  - e) Wyższe
  
4. Sytuacja zawodowa
  - a) Uczę się
  - b) Bezrobotny/-a
  - c) Pracujący/-a
  - d) Emeryt/-ka / Rencista/-ka

Warunkowe przejście do kolejnej sekcji:

Odpowiedź a) – przejdź do sekcji „Osoby niepracujące”

Odpowiedź b) – przejdź do sekcji „Sytuacja zawodowa w przeszłości”

Odpowiedź c) – przejdź do sekcji „Osoby pracujące (obecnie lub w przeszłości)”

Odpowiedź d) – przejdź do sekcji „Sytuacja zawodowa w przeszłości”

### **Sekcja: Sytuacja zawodowa w przeszłości**

5. Czy w przeszłości był/-a Pan/Pani aktywny/-a zawodowo?

- a) Tak
- b) Nie

Warunkowe przejście do kolejnej sekcji:

Odpowiedź a) – przejdź do sekcji „Osoby pracujące (obecnie lub w przeszłości)”

Odpowiedź b) – przejdź do sekcji „Osoby niepracujące”

### **Sekcja: Osoby pracujące (obecnie lub w przeszłości)**

6. Jakiej branży dotyczy Pana/Pani aktywność zawodowa?

- a) Usługi
- b) Przedsiębiorstwo produkcyjne
- c) Administracja
- d) Budownictwo
- e) IT
- f) Gastronomia
- g) Inne: ...

7. Jestem osobą:

- a) Niedosłyszającą
- b) Niesłyszającą
- c) Słyszającą

Warunkowe przejście do kolejnej sekcji:

Odpowiedź a) – przejdź do sekcji „Osoby niesłyszące i niedosłyszające”

Odpowiedź b) – przejdź do sekcji „Osoby niesłyszące i niedosłyszające”

Odpowiedź c) – przejdź do sekcji „Osoby słyszące lub nieznaną języka migowego”

### **Sekcja: Osoby niesłyszące i niedosłyszające**

8. Z jakimi trudnościami spotkał/-a się Pan/Pani w wykonywanej przez Pana/Panią pracy?

(1 - brak wystąpień, 5 - częste występowanie)

- a) Komunikacja z drugą osobą ...
- b) Dyskryminacja ze względu na niepełnosprawność ...
- c) Możliwość rozwoju ...
- d) Niemożność wykonania polecenia ...



e) Niezrozumienie (kogoś lub przez kogoś) ...

9. Proszę zaznaczyć technologie/urządzenia, które wykorzystywał/-a Pan/Pani przy wykonywaniu obowiązków zawodowych. *(Pytanie wielokrotnego wyboru)*

- a) Aparat słuchowy
- b) Syntezator mowy (przekształcanie tekstu na mowę)
- c) Analizator mowy (przekształcanie mowy na tekst)
- d) Nie korzystałem/-am z żadnych technologii/urządzeń

10. Czy zna Pan/Pani język migowy?

- a) Tak
- b) Nie

Warunkowe przejście do kolejnej sekcji:

Odpowiedź a) – przejdź do sekcji „Osoby znające język migowy”

Odpowiedź b) – przejdź do sekcji „Osoby słyszące lub nieznające języka migowego”

### **Sekcja: Osoby znające język migowy**

11. W jakim stopniu zna Pan/Pani język migowy? (1 - w ogóle, 5 - biegle)

- a) PJM ...
- b) ASL ...
- c) BSL ...
- d) GESTUNO ...

12. Czy Pana/Pani zdaniem wykorzystanie rękawic tłumaczących język migowy na tekst i mowę ułatwiłoby osobom z niepełnosprawnościami słuchowymi komunikację w pracy?

- a) Tak
- b) Nie

13. W jakim stopniu postrzega Pan/Pani potencjalne problemy związane ze stosowaniem inteligentnych rękawic? (1 - nie widzę problemu, 5 - poważny problem)

- a) Spowolniona prędkość wykonywanych gestów ...
- b) Czas pracy na baterii ...
- c) Waga urządzenia ...

- d) Skuteczność rozpoznawania gestów ...
- e) Ograniczony zasób słów ...
- f) Konieczność stosowania aplikacji mobilnej ...
- g) Cena urządzenia ...

14. Jakie usprawnienia można by wdrożyć w projekcie inteligentnych rękawic? *(Pytanie otwarte)*

### **Koniec ankiety**

### **Sekcja: Osoby słyszące lub nieznające języka migowego**

15. Czy pracował/-a Pan/Pani z osobą z niepełnosprawnością słuchową?

- a) Tak
- b) Nie

Warunkowe przejście do kolejnej sekcji:

Odpowiedź a) – przejdź do sekcji „Osoby współpracujące z osobami z niepełnosprawnością słuchową”

Odpowiedź b) – przejdź do sekcji „Osoby słyszące lub nieznające języka migowego (ciąg dalszy)”

### **Sekcja: Osoby współpracujące z osobami z niepełnosprawnością słuchową**

16. Z jakimi trudnościami spotykała się osoba z niepełnosprawnością słuchową w wykonywanej przez nią pracy? (1 - brak wystąpień, 5 - częste występowanie)

- a) Komunikacja z drugą osobą ...
- b) Dyskryminacja ze względu na niepełnosprawność ...
- c) Możliwość rozwoju ...
- d) Niemożność wykonania polecenia ...
- e) Niezrozumienie (kogoś lub przez kogoś) ...

17. Proszę zaznaczyć technologie/urządzenia, które wykorzystywała osoba z niepełnosprawnością słuchową przy wykonywaniu obowiązków zawodowych.

- a) Aparat słuchowy
- b) Syntezator mowy (przekształcanie tekstu na mowę)
- c) Analizator mowy (przekształcanie mowy na tekst)
- d) Nie korzystałem/-am z żadnych technologii/urządzeń

**Sekcja: Osoby słyszące lub nieznające języka migowego (ciąg dalszy)**

18. Czy Pana/Pani zdaniem wykorzystanie rękawic tłumaczących język migowy na tekst i mowę ułatwiłoby osobom z niepełnosprawnościami słuchowymi komunikację w pracy?

- a) Tak
- b) Nie

19. W jakim stopniu postrzega Pan/Pani potencjalne problemy związane ze stosowaniem inteligentnych rękawic? (1 - nie widzę problemu, 5 - poważny problem)

- a) Spowolniona prędkość wykonywanych gestów ...
- b) Czas pracy na baterii ...
- c) Waga urządzenia ...
- d) Skuteczność rozpoznawania gestów ...
- e) Ograniczony zasób słów ...
- f) Konieczność stosowania aplikacji mobilnej ...
- g) Cena urządzenia ...

20. Jakie usprawnienia można by wdrożyć w projekcie inteligentnych rękawic? (*Pytanie otwarte*)

**Koniec ankiety**

**Sekcja: Osoby niepracujące**

21. Czy Pana/Pani zdaniem wykorzystanie rękawic tłumaczących język migowy na tekst i mowę ułatwiłoby osobom z niepełnosprawnościami słuchowymi komunikację w pracy?

- a) Tak
- b) Nie

22. W jakim stopniu postrzega Pan/Pani potencjalne problemy związane ze stosowaniem inteligentnych rękawic? (1 - nie widzę problemu, 5 - poważny problem)

- a) Spowolniona prędkość wykonywanych gestów ...
- b) Czas pracy na baterii ...
- c) Waga urządzenia ...

- d) Skuteczność rozpoznawania gestów ...
- e) Ograniczony zasób słów ...
- f) Konieczność stosowania aplikacji mobilnej ...
- g) Cena urządzenia ...

23. Jakie usprawnienia można by wdrożyć w projekcie inteligentnych rękawic? (*Pytanie otwarte*)

**Koniec ankiety**

## **Załącznik 3. Spis zawartości płyty CD dołączonej do pracy**

1. Praca magisterska w formacie \*.doc i \*.pdf.