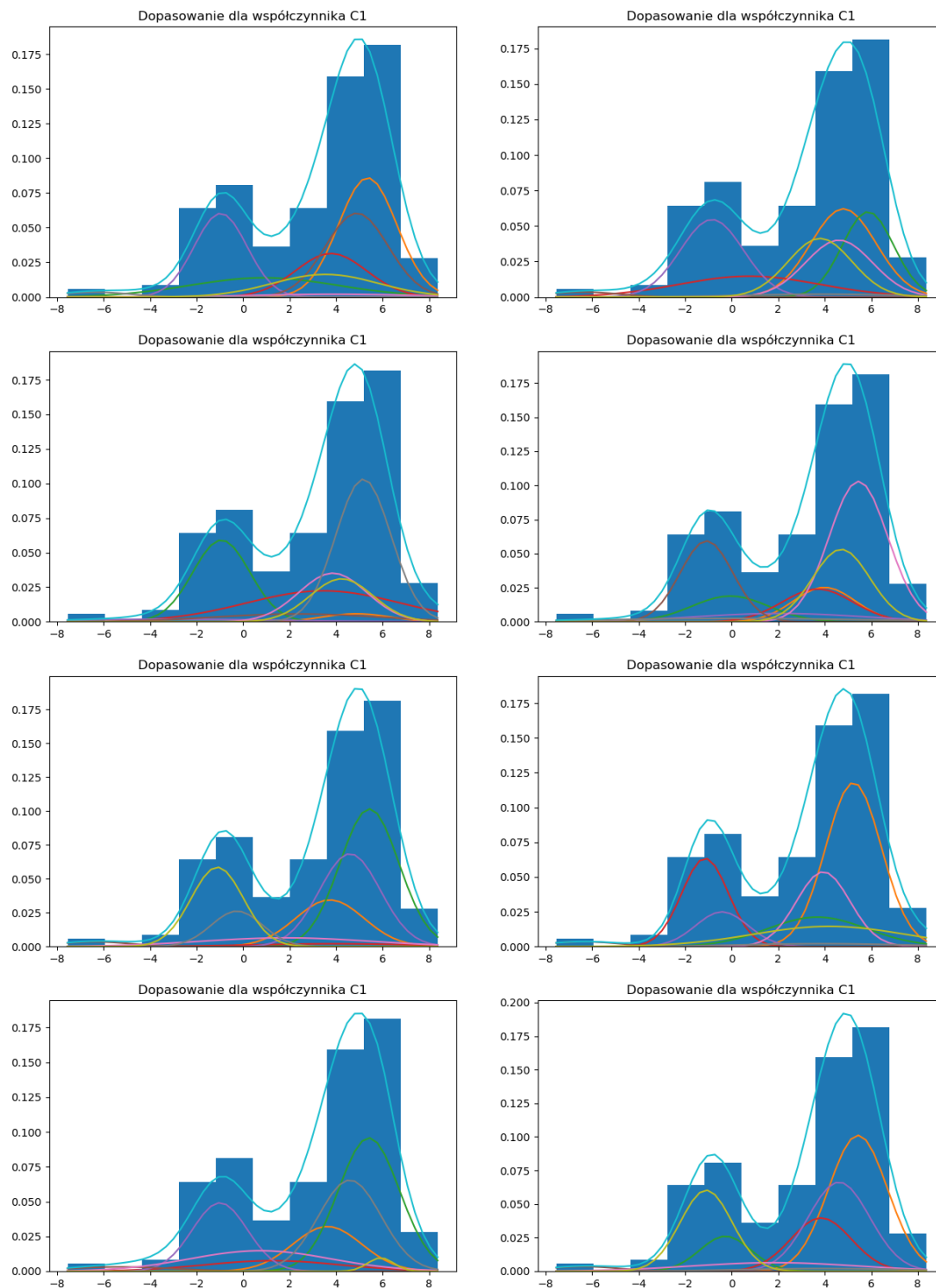


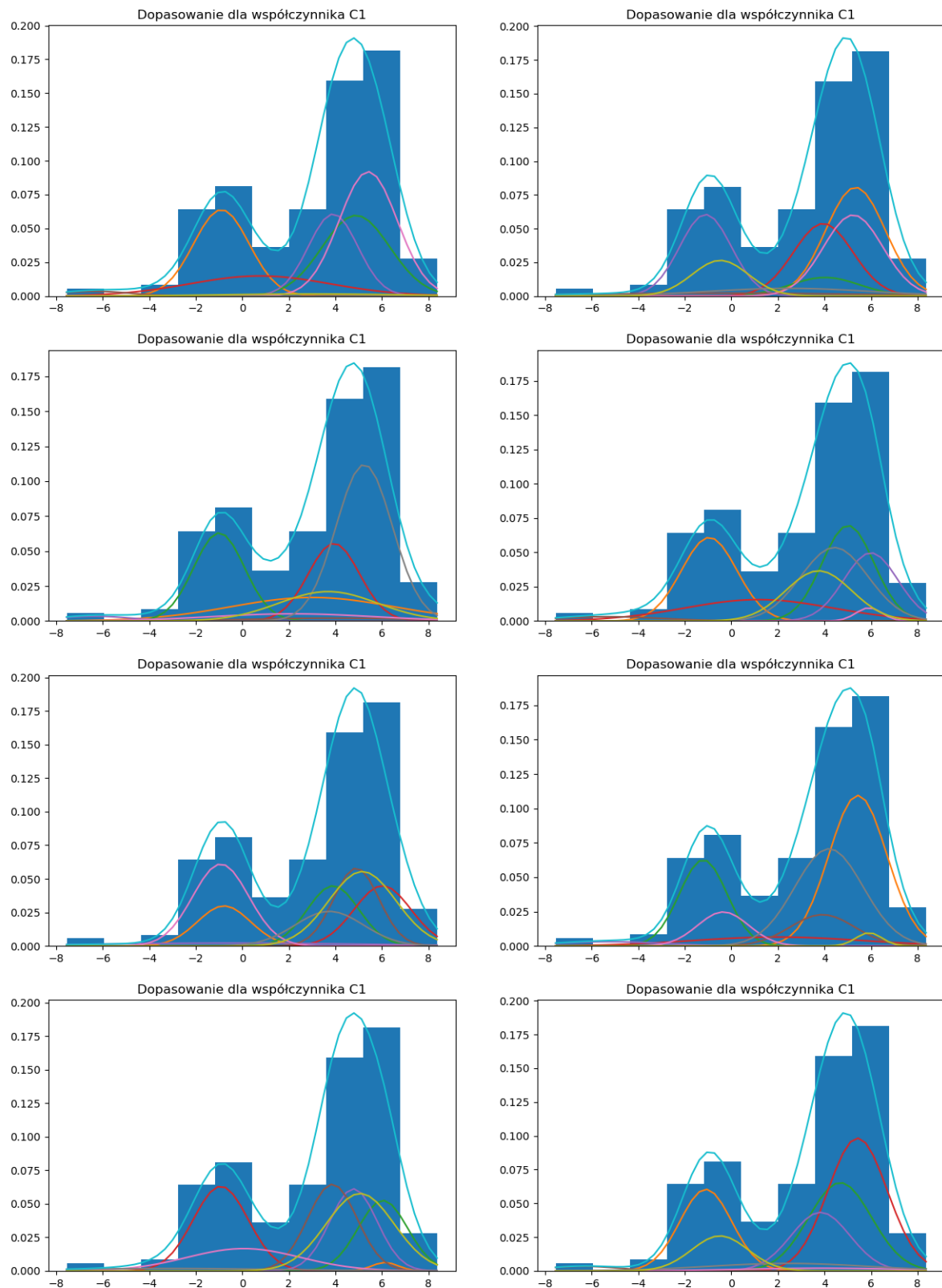
## ZADANIE 1 – MODEL GMM

## WYKRESY DLA RÓŻNEJ LICZBY ITERACJI ALGORYTMU EM



Rysunek 1. Dopasowania dla różnej ilości iteracji algorytmu EM (1-8)

## TECHNOLOGIA MOWY – RAPORT 4



Rysunek 2. Dopasowania dla różnej wartości iteracji algorytmu EM (10-80 w kroku co 10)

## TECHNOLOGIA MOWY – RAPORT 4

Analizując wykresy dla różnej wartości iteracji algorytmu EM można dostrzec iż dla niższych wartości dopasowania nie zawsze pasują do histogramu. Jest to kwestia w pełni losowa czy dopasowane krzywe będą dobrze odzwierciedlać daną cechę, czy też będą nie dość dobre (tak jak widać wyżej na wykresie dla 2 i 7 współczynników). Im większa liczba iteracji tym wykresy robią się coraz bardziej podobne do siebie, oraz do wartości dla konkretnego współczynnika.

## ANALIZA WARTOŚCI KRYTERIUM INFORMACYJNEGO W ZALEŻNOŚCI OD LICZBY ITERACJI

Tabela 1. Wartości kryterium informacyjnego w zależności od liczby iteracji algorytmu EM

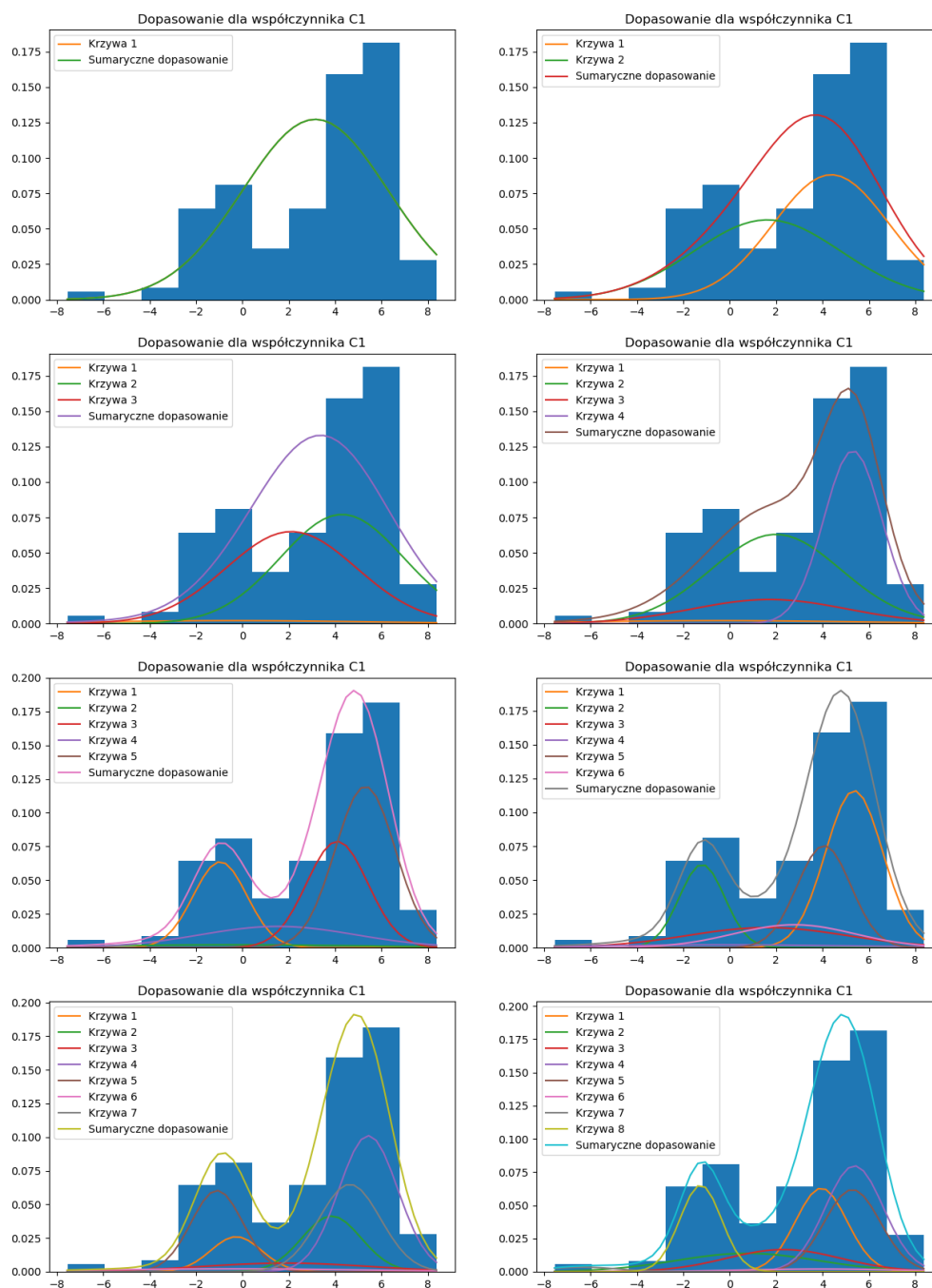
Liczba iteracji	AIC	BIC
1	13599	14334
3	13576	14311
5	13449	14183
7	13476	14210
9	13421	14155
10	13528	14262
20	13404	14138
30	13491	14226
50	13497	14232
70	13445	14179
90	13492	14227

Stwierdzam, iż jako kryterium przyjmuję AIC, gdyż różnice w wartości parametru AIC większe niż 10 gwarantują iż dopasowanie modelu będzie dobre. Poza tym różnice w wartości parametru BIC są zdecydowanie mniejsze niż dla AIC. Dodatkowo BIC jest parametrem, który z definicji radzi sobie lepiej w opisywaniu modeli prostych.

## ILE KOMPONENTÓW JEST OPTIMALNYCH DLA MODELOWANIA GŁOSEK

## TECHNOLOGIA MOWY – RAPORT 4

Zgodnie z powyższymi wykresami, można dostrzec, iż liczba komponentów wystarczających do poprawnego



zamodelowania głosek to 6. Każda mniejsza wartość sprawia iż sumaryczne dopasowanie nie jest na tyle poprawnym aby dobrze zamodelować współczynniki (w powyższym przypadku współczynnik C1).

## ZADANIE 2 – LOG-PRAWDOPODOBIENSTWA

## OBLICZONE WARTOŚCI LOG-PRAWDOPODOBIENSTW

Tabela 2. Wartości log-prawdopodobieństw

$\log(p(MFCC_A GMM_A))$	-28,99
$\log(p(MFCC_I GMM_A))$	-257,22
$\log(p(MFCC_I GMM_I))$	-29,93
$\log(p(MFCC_A GMM_I))$	-336,76

Uzyskane wartości są zgodne z oczekiwaniami. Pokazują iż nagranie głoski /a/ pasuje do modelu dla tej głoski, a nagranie głoski /i/ pasuje do modelu głoski /i/. Wartości prawdopodobieństw odwrotnych (jak nagranie głoski /a/ pasuje do modelu głoski /i/ oraz na odwrót) są dużo niższe, co wskazuje na to że dopasowanie głoski /a/ do modelu /i/ oraz na odwrót jest bardzo mało prawdopodobne.

## WARTOŚCI PRAWDOPODOBIENSTW Z REGUŁY BAYESA DLA NOWO UTWORZONYCH NAGRAŃ

$$\log(p(MFCC_A|GMM_A)) = -29.08$$

$$\log(p(MFCC_I|GMM_A)) = -386.01$$

$$\log(p(MFCC_I|GMM_I)) = -30.13$$

$$\log(p(MFCC_A|GMM_I)) = -569.66$$

$$p(GMM_A|MFCC_A) = 0.95$$

$$p(GMM_A|MFCC_I) = 9.7 * 10^{-156}$$

$$p(GMM_I|MFCC_A) = 4.8 * 10^{-235}$$

$$p(GMM_I|MFCC_I) = 0.89$$

Nowe nagrania uzyskują bardzo wysokie prawdopodobieństwo zaklasyfikowania jako odpowiednia głoska. Dla próby przypasowania do niepasującego modelu prawdopodobieństwo jest tak marginalnie małe, że można uznać iż wynosi zero.