Analiza Big Data informacji pochodzących z urządzeń końcowych STB oraz wykorzystanie metod uczenia maszynowego w celu usprawnieniu działania usługi, analizie doświadczeń klienta oraz zaspokojeniu jego oczekiwań

Big data analysis of information from STB end devices and the use of machine learning methods to improve the operation of the service with the customer experience analysis

Paweł Andruloniw

Wydział Informatyki i Telekomunikacji
Politechnika Poznańska

Rozprawa doktorska

Abstract
(Streszczenie)

Promotorzy:
dr hab. inż. Piotr Zwierzykowski, prof. PUT
dr inż. Karol Kowalik

Poznań 2025

Abstract

The development of television systems and methods for delivering video content to end users poses new challenges for telecommunications operators. These entities are required to continuously modernize their headend infrastructure in order to meet growing user expectations and ensure an appropriate level of service quality. HTTP adaptive streaming constitutes one of the key methods of video content distribution, enabling the provision of television services both in the form of linear channels and non-linear, time-shifted services. One of the fundamental aspects of delivering a reliable television service is maintaining an adequate level of quality of the distributed content. Consequently, increasingly advanced methods are being developed to assess the quality of video delivered to end users. The subjective impressions of viewers regarding perceived image quality arising from the characteristics of the human visual system form the basis of methods used to evaluate perceived quality, known as *QoE* (ang. *Quality of Experience*).

The assessment of user experience quality (QoE) considers not only objective signal parameters, such as image quality determined by the encoding standard, stream bitrate, bitrate variability, or the presence of playback stalls, but also the subjective impressions of viewers using the television service. Obtaining such an assessment requires the application of standardized methods for collecting subjective quality scores. In the QoE analysis presented in this work, the *DSCQS* (ang. *Double-Stimulus Continuous Quality Scale*) method was employed to determine the subjective perception of video quality. The study was conducted for the Hiway television system operated by INEA. As part of the experiment, two test sequences were prepared, each with a duration of slightly more than 23 minutes. Each sequence consisted of 30 video materials organized into 15 reference – impairment pairs, and each material was independently evaluated with respect to QoE.

The study was designed to allow end users to participate directly through their set-top boxes, thereby enabling the conditions of real-world television service usage to be reproduced as faithfully as possible. To collect subjective assessments, a dedicated mobile application was developed to aggregate feedback from the participants. Each Hiway television subscriber was able to comfortably sit in front of their television and complete the assessment in the privacy of their home. The DSCQS methodology also enables group evaluations, which made it possible to reflect typical conditions of watching television in a larger audience. In total, 3,240 ratings were collected from 108 participants for video materials encoded using the H.264 and H.265 standards, prepared across a bitrate range from 800 kbps to 10,000 kbps.

The results obtained in the study served as the basis for the development of an algorithm employing an artificial neural network. A multilayer perceptron (MLP) architecture was used to construct a mechanism for predicting QoE scores for Hiway service users. The estimation task was formulated as a multi-class classification problem with soft labels, in which the predicted output represents a probability distribution over the possible QoE ratings, conditioned on the provided input parameters. These parameters encompassed both the characteristics of the video materials and the subjective assessments collected from participants in accordance with the DSCQS procedure.

Figure 1 presents the general architecture of the neural network designed for predicting ratings for Hiway TV users. The modeling process was conducted in three stages. In the first step, a baseline

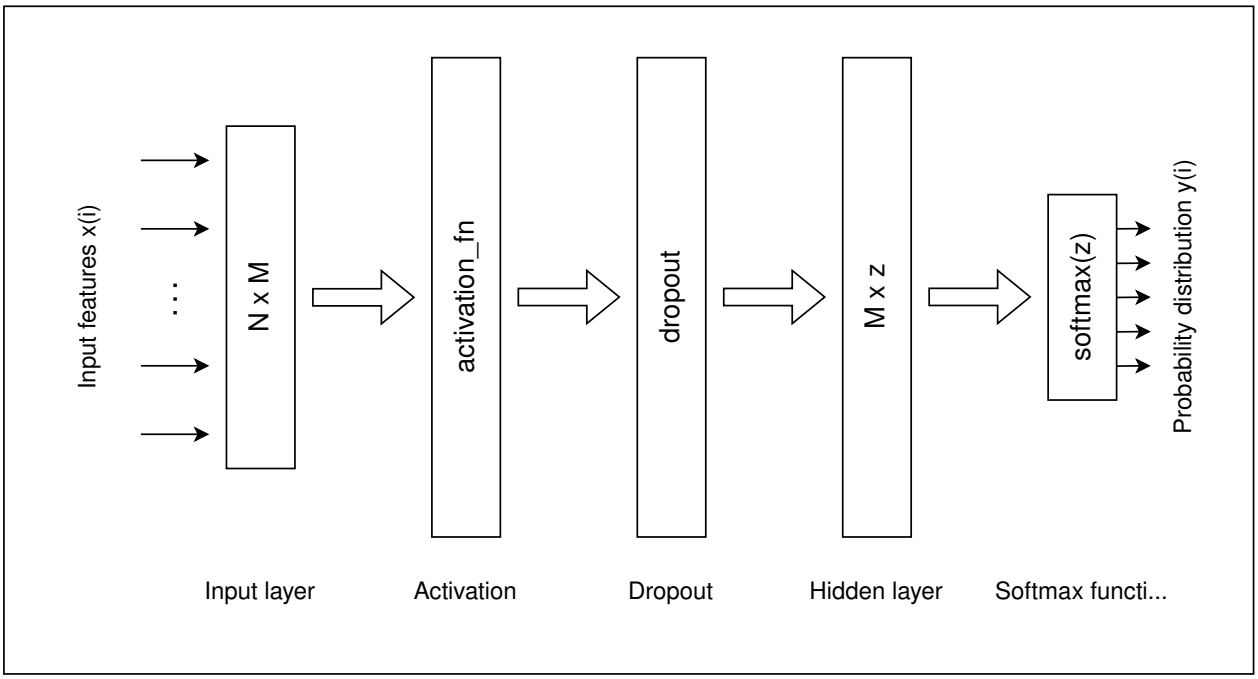


Figure 1 Baseline neural network model used for multi-class classification with soft labels

model was developed and trained, serving as a benchmark for subsequent work. The baseline neural network model consisted of a linear input layer and a hidden layer, which was a combination of a linear layer, an activation function, and a regularization layer. To enhance model generalization, the dropout regularization technique was employed, enabling the deactivation of selected neurons along with their inputs and outputs. The linear output layer possessed M inputs and Z outputs, where $Z = 5$. The output of this layer returns a five-element vector of unnormalized numerical values (logits). The softmax activation function converts this unnormalized input vector into a five-element probability vector. The sum of these values is 1, and each value represents the probability of a specific rating on the ACR (Absolute Category Rating) scale for the given input features. Three baseline models were trained, which served to indicate the direction of further optimization work.

Subsequently, a two-stage optimization was conducted on successive neural network architectures, analyzing their effectiveness in predicting the Quality of Experience (QoE) metric. The optimization was focused on maximizing the Spearman's Rank Correlation Coefficient (SROCC), as well as key classification metrics, including Accuracy, Precision, Recall, and the F1 measure. In the first phase, the focus was on shallow architectures, characterized by a single hidden layer with a variable number of neurons. During this optimization, the influence of the batch size, the non-linear activation function employed, and the values of key hyperparameters (e.g., learning rate) on the final results were verified. The second phase involved analyzing deep models, which possessed more than one hidden layer. In total, over 400 diverse neural network models were trained. The most effective models were then re-verified using 10-fold cross-validation to confirm their stability and generalization capabilities.

Ultimately, three models were selected, characterized by the highest consistency with the subjective ratings provided by the QoE study participants. Figure 2 presents the prediction results for the QoE metric across four exemplary samples, each corresponding to a 60-second video sequence. This

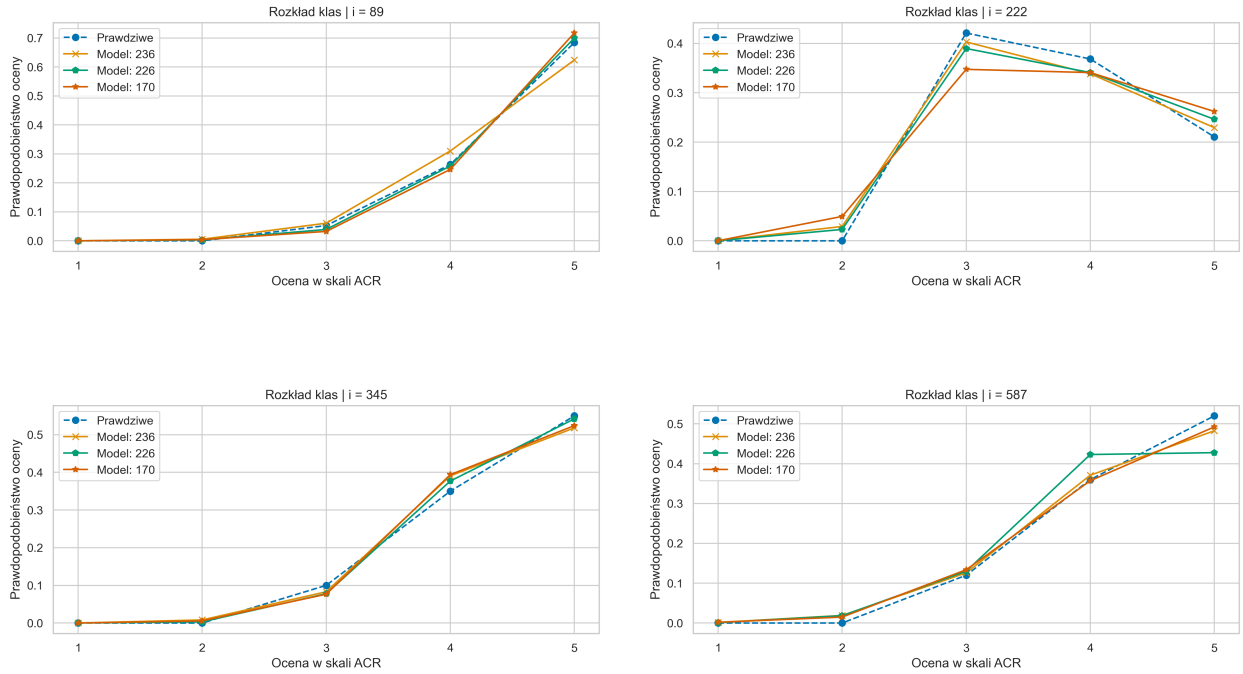


Figure 2 Estimation of the probability distribution for exemplary samples by the three most effective models

sequence mirrors the QoE assessment for two 30-second video segments, aligning with the approach previously employed in the DSCQS (Double-Stimulus Continuous Quality Scale) methodology.The prediction results are presented as the probability distribution across the five rating categories on the ACR scale (1 to 5) for each 60-second sample. Four data series are displayed on each plot. One series represents the averaged probability distribution of the subjective scores given by the study participants (the reference data). The remaining three series represent the estimation results obtained by the selected soft classification models. For the most effective models, the average values of the Spearman's Rank Correlation Coefficient (SROCC), Accuracy, and the F1 measure achieved a level of above 93%.

An additional, practical implementation aspect of this work is the design of a tool, based on the developed QoE estimator, intended for customer service departments. This tool, presented in Figure 3, enables the predictive assessment of Quality of Experience (QoE) for users utilizing the Hiway TV service. Analyzing the QoE level for selected set-top boxes allows for the reconstruction of the user's perception of quality during a specific broadcast, channel, or entire viewing session (i.e., from device startup to shutdown). Consequently, the proposed solution offers two key benefits: firstly, it provides a better understanding of the subjective feelings of customers reporting issues, and secondly, it enables the proactive detection of devices characterized by a reduced QoE level, which in turn allows for faster corrective actions and interventions before the service user formally reports the problem.

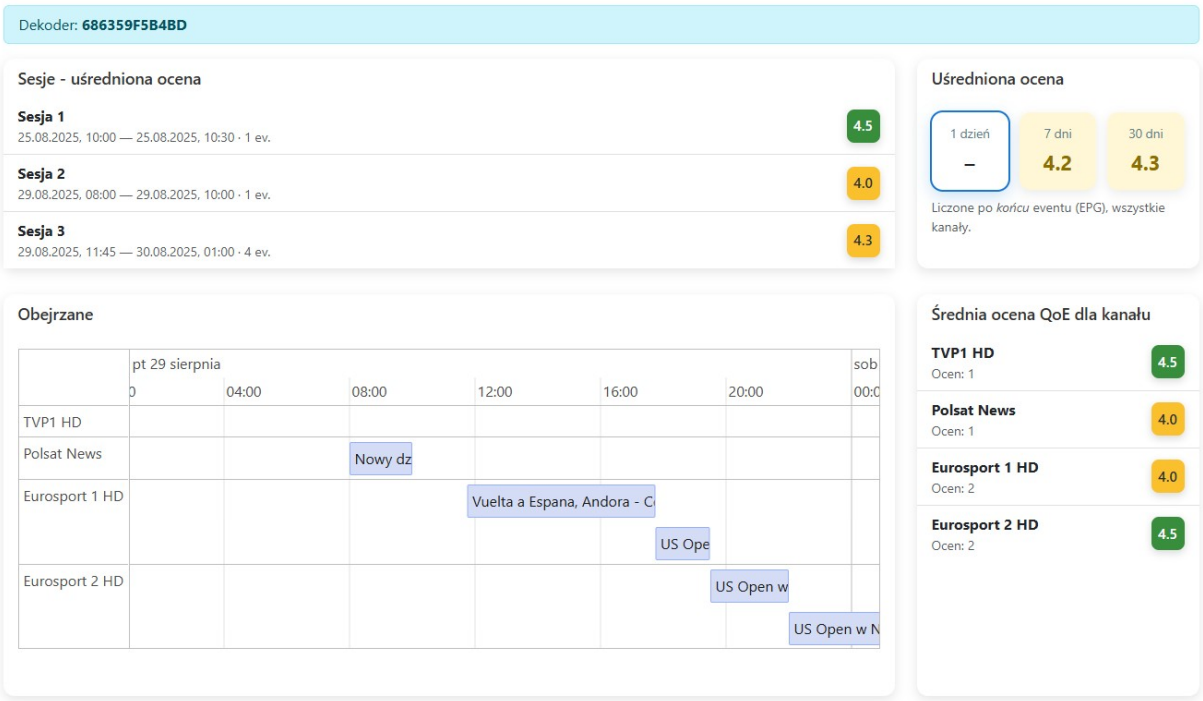


Figure 3 A tool for assessing QoE in the television service with an implemented QoE estimator based on subjective quality assessment

Publications of the author

Papers in scored journals

1. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Subiektywne postrzeganie pojedynczych zniekształceń w telewizji opartej o strumieniowanie adaptacyjne*, *Przegląd Telekomunikacyjny – Wiadomości Telekomunikacyjne*, 2024, nr 4, s. 249–252.

2. P. Andruloniw, K. Kowalik, B. Partyka, P. Zwierzykowski: *Multi-layer QoE Learning System Implemented by Fiberhost*, *Applied Sciences*, 2023, vol. 13, iss. 4, s. 2300-1–2300-23.

3. K. Kowalik, P. Andruloniw, B. Partyka, P. Zwierzykowski: *Telecom Operator's Approach to QoE*, *Journal of Telecommunications and Information Technology*, 2022, nr 2, s. 26–34.

4. P. Andruloniw, K. Kowalik, P. Zwierzykowski: *Unsupervised Learning Data-Driven Continuous QoE Assessment in Adaptive Streaming-Based Television System*, *Applied Sciences*, 2022, vol. 12, iss. 16, s. 8288-1–8288-23.

Book chapters and conference papers

1. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Quality of Experience Studies in TV Systems: Insights and Lessons Learned*, w: *Artificial Intelligence and Machine Learning. 43rd IBIMA Conference, Madrid, 2024*, Springer Nature, 2025, s. 267–273.

2. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Subiektywna ocena jakości wideo w sieci dostawcy dla usługi telewizyjnej opartej o strumieniowanie adaptacyjne*, w: *Perspektywy rozwoju w naukach inżynieryjno-technicznych – trendy, innowacje i wyzwania. Tom 1*, Wyd. Naukowe TYGIEL, 2024, s. 104–114.

3. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Quality of Experience Data Gathering in OTT Television Service*, w: *Proceedings of the 41st IBIMA Conference, Granada, 2023*, IBIMA Publishing, 2023, s. 119–122.

4. P. Andruloniw, K. Kowalik, P. Zwierzykowski: *Analiza skupień dla skończonego oraz ciągłego strumienia danych*, w: *Rozwiązania technologiczne XXI wieku – skutki i perspektywy rozwoju. Tom 1*, Wyd. Naukowe TYGIEL, 2022, s. 7–22.

5. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Analysis of Adaptive Video Streaming Degradation Based on Clustering Approach*, w: *Proceedings of the 39th IBIMA Conference, Granada, 2022*, IBIMA Publishing, 2022, s. 192–199.

6. P. Andruloniw, P. Zwierzykowski: *W jaki sposób metody uczenia maszynowego mogą pomóc w wykrywaniu degradacji i szacowaniu jakości usług wideo w sieciach telekomunikacyjnych?*, w: *Technologie XXI wieku – aktualne problemy i nowe wyzwania. Tom 3*, Wyd. Naukowe TYGIEL, 2021, s. 132–141.

Sum of MEiN points: 350, **Sum of CORE points: 530**

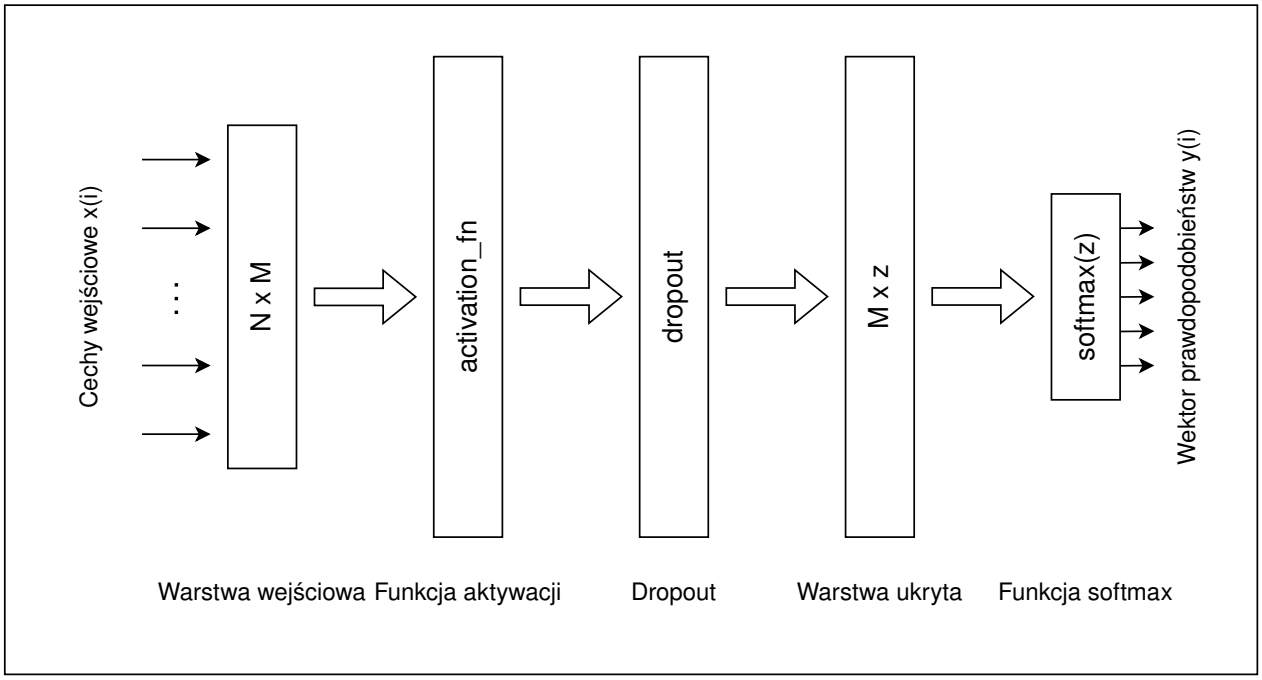
Streszczenie

Rozwój systemów telewizyjnych oraz metod dystrybucji treści wideo do odbiorców końcowych generuje nowe wyzwania dla operatorów telekomunikacyjnych. Podmioty te są zmuszone do ciągłej modernizacji infrastruktury stacji czołowej, aby sprostać rosnącym wymaganiom użytkowników oraz zapewnić odpowiedni poziom jakości usług. Strumieniowanie adaptacyjne HTTP stanowi jedną z kluczowych metod dystrybucji treści wideo, umożliwiającą realizację usług telewizyjnych zarówno w formie kanałów liniowych, jak i nieliniowych usług dodatkowych związanych z przesuwaniem treści w czasie ang. *timeshift*. Jednym z kluczowych aspektów świadczenia niezawodnej usługi telewizyjnej jest zapewnienie odpowiedniego poziomu jakości dystrybuowanych treści. W związku z tym opracowywane są coraz bardziej zaawansowane metody oceny jakości wideo dostarczanego do odbiorcy końcowego. Subiektywne wrażenia użytkowników dotyczące postrzeganej jakości obrazu, wynikające z właściwości ludzkiego systemu wzrokowego, stanowią podstawę metod oceny jakości postrzeganej *QoE* (ang. *Quality of Experience*).

Ocena jakości doświadczeń użytkownika (QoE) uwzględnia nie tylko obiektywne parametry sygnału — takie jak jakość obrazu wynikająca ze standardu kodowania, przepływność bitowa strumienia, jej zmienność czy występowanie zacięć — lecz również subiektywne odczucia odbiorców korzystających z usługi telewizyjnej. Uzyskanie takiej oceny wymaga zastosowania standaryzowanych metod pozyskiwania subiektywnych ocen jakości. W analizie QoE zaprezentowanej w niniejszej pracy wykorzystano metodę *DSCQS* (ang. *Double-Stimulus Continuous Quality Scale*), umożliwiającą określenie subiektywnego postrzegania jakości treści wideo. Badanie przeprowadzono dla systemu telewizyjnego Hiway, należącego do operatora INEA. W ramach eksperymentu przygotowano dwie sekwencje badawcze, trwające po nieco ponad 23 minuty każda. Każda z nich obejmowała 30 materiałów filmowych zorganizowanych w 15 par typu referencja–bodziec, z których każdy podlegał niezależnej ocenie w kontekście QoE.

Badanie zostało zaprojektowane w sposób umożliwiający jego realizację przez użytkowników końcowych bezpośrednio za pośrednictwem dekoderów, co pozwoliło na możliwie wierne odwzorowanie rzeczywistych warunków korzystania z usługi telewizyjnej. W celu pozyskania ocen subiektywnych opracowano mobilną aplikację wspierającą proces agregacji informacji zwrotnej od uczestników eksperymentu. Każdy użytkownik telewizji Hiway mógł usiąść przed telewizorem i wykonać badanie w domowym zaciszu. Formuła badania DSCQS umożliwia także jego przeprowadzenie w grupie, co pozwoliło na odwzorowanie warunków wspólnego oglądania telewizji. Łącznie zebrano 3240 ocen od 108 osób dla materiałów zakodowanych zgodnie ze standardami H.264 oraz H.265, przygotowanych w zakresach przepływności bitowej od 800 kbit/s do 10 000 kbit/s.

Uzyskane wyniki stanowiły podstawę do opracowania algorytmu wykorzystującego sztuczną sieć neuronową. Zastosowano architekturę perceptronu wielowarstwowego (MLP) w celu zbudowania mechanizmu predykcji ocen QoE dla użytkowników usługi Hiway. Problem szacowania ocen sprowadzono do klasyfikacji wieloklasowej z wykorzystaniem miękkich etykiet, gdzie estymowaną wartość reprezentuje rozkład prawdopodobieństwa ocen QoE na podstawie zadanych parametrów wejściowych. Parametry te obejmowały zarówno cechy materiałów wideo, jak i subiektywne oceny pozyskane od uczestników badania zgodnie z procedurą DSCQS.



Rysunek 1 Bazowy model sieci neuronowej wykorzystany w klasyfikacji wieloklasowej z miękkimi etykietami

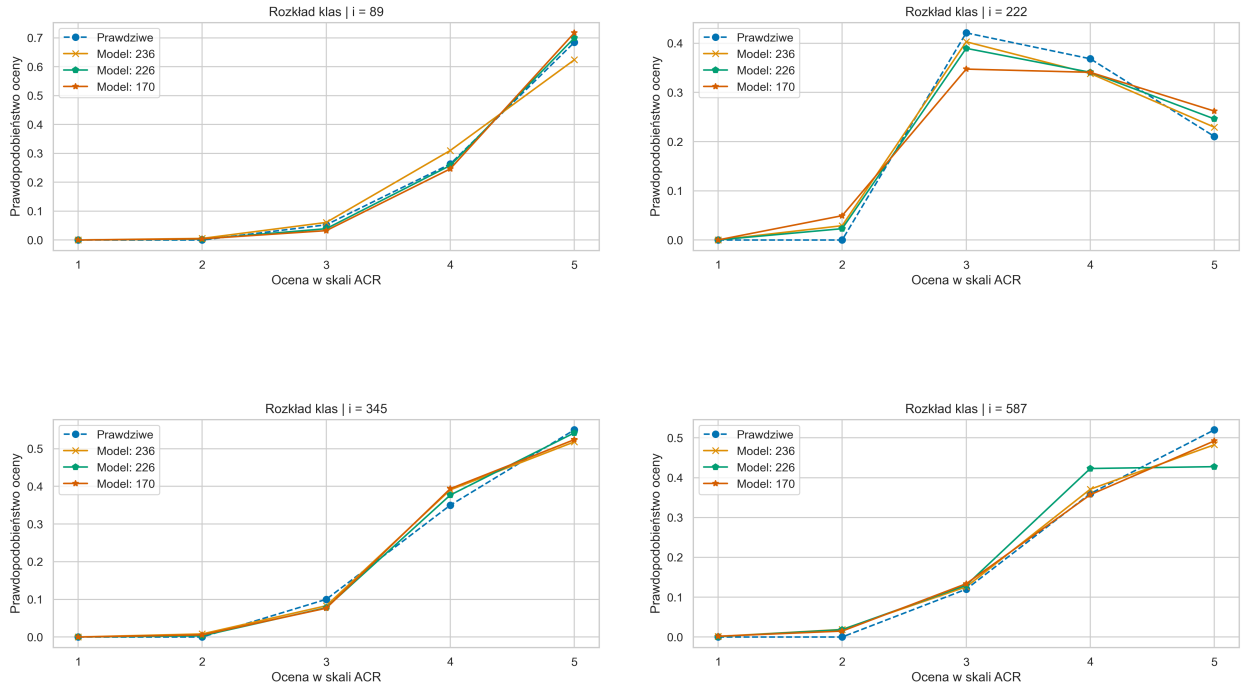
Na rysunku 1 przedstawiono ogólną architekturę sieci neuronowej do predykcji ocen dla użytkowników telewizji Hiway. Proces modelowania przeprowadzono w trzech etapach. W pierwszym kroku opracowano i wytrenowano model bazowy, stanowiący punkt odniesienia do dalszych prac. Bazowy model sieci neuronowej składał się z liniowej warstwy wejściowej oraz warstwy ukrytej będącej kombinacją warstwy liniowej, aktywacji oraz warstwy regularyzacyjnej. Do generalizacji modelu zastosowano technikę regularyzacji opuszczeń (ang. *dropout*) umożliwiającą dezaktywację wybranych neuronów wraz z ich wejściami i wyjściem. Liniowa warstwa wyjściowa posiada M wejść oraz Z wyjść, gdzie $Z = 5$. Na wyjściu tej warstwy zwracany jest pięcioelementowy wektor nieznormalizowanych wartości liczbowych (ang. *logits*). Funkcja aktywacji *softmax* przekształca nieznormalizowany wektor wejściowy w pięcioelementowy wektor prawdopodobieństw. Suma wartości wynosi 1, a każda wartość to prawdopodobieństwo danej oceny w skali ACR dla zadanych cech wejściowych. Wytrenowano trzy modele bazowe, które wskazały kierunek dalszych prac.

Następnie przeprowadzono dwuetapową optymalizację kolejnych architektur sieci neuronowych, analizując ich skuteczność w prognozowaniu wskaźnika jakości doświadczenia (QoE). Optymalizację ukierunkowano na maksymalizację współczynnika korelacji rang Spearmana (SROCC), a także kluczowych metryk klasyfikacji, w tym celności (Accuracy), precyzji (Precision), czułości (Recall) oraz miary F1. W pierwszej fazie skupiono się na architekturach płytkich, charakteryzujących się jedną warstwą ukrytą o zmiennej liczbie neuronów. W trakcie optymalizacji weryfikowano wpływ liczby próbek przetwarzanych jednocześnie ang. *(*batch size), stosowanej nieliniowej funkcji aktywacji oraz wartości kluczowych hiperparametrów (np. współczynnika uczenia) na ostateczne wyniki. W drugiej fazie analizowano modele głębokie, posiadające więcej niż jedną warstwę ukrytą. Sumarycznie wytrenowano ponad 400 zróżnicowanych modeli sieci neuronowych. Najskuteczniejsze

modele były następnie ponownie weryfikowane przy użyciu 10-krotnej walidacji krzyżowej w celu potwierdzenia ich stabilności i generalizacji.

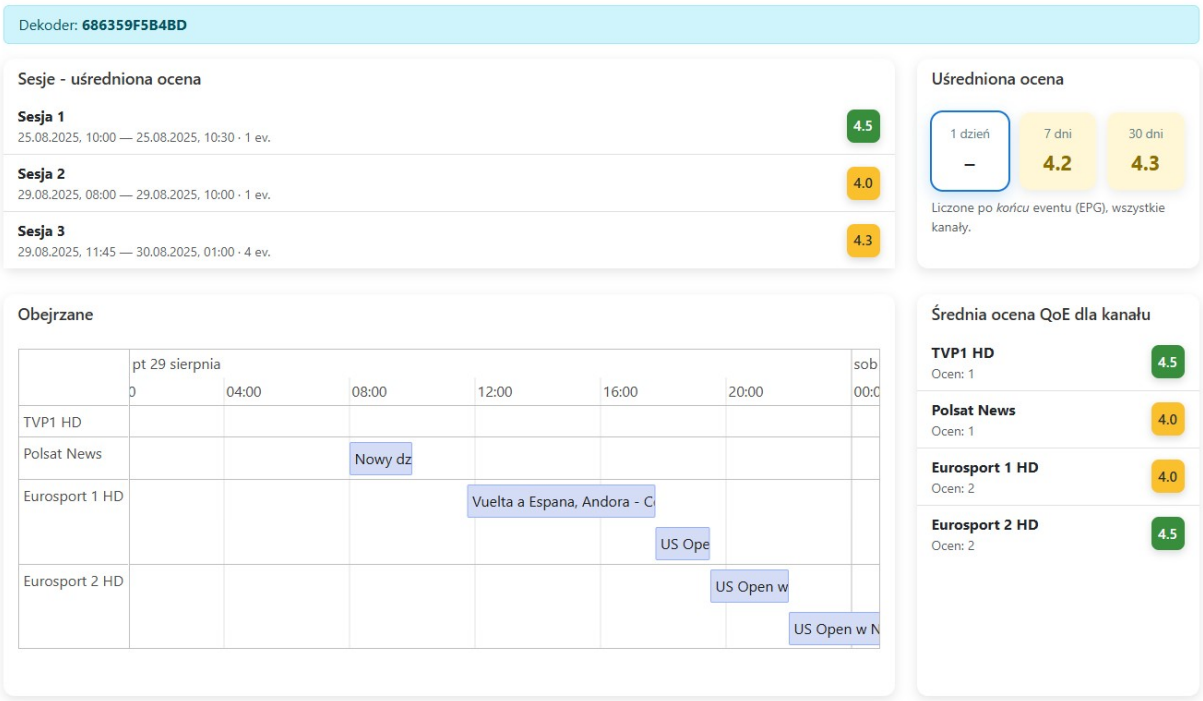
Ostatecznie, wybrano trzy modele charakteryzujące się najwyższą zgodnością z subiektywnymi ocenami uczestników badania QoE. Rysunek 2 przedstawia wyniki predykcji jakości QoE dla czterech przykładowych próbek, z których każda odpowiada 60-sekundowej sekwencji wideo. Sekwencja ta odwzorowuje szacowanie QoE dla dwóch 30-sekundowych materiałów wideo, zgodnie z podejściem stosowanym wcześniej w badaniu DSCQS (ang. *Double-Stimulus Continuous Quality Scale*). Wyniki predykcji przedstawiono jako rozkład prawdopodobieństwa pięciu kategorii ocen w skali ACR (od 1 do 5) dla każdej 60-sekundowej próbki. Na każdym wykresie zaprezentowano cztery serie danych. Jedna seria stanowi uśredniony rozkład prawdopodobieństwa subiektywnych ocen udzielonych przez uczestników badania (dane referencyjne). Pozostałe trzy serie reprezentują wyniki estymacji uzyskane przez wybrane modele klasyfikacji miękkiej. Dla najskuteczniejszych modeli, średnie wartości współczynnika korelacji rang Spearmana (SROCC), celności (Accuracy) oraz miary F1 osiągnęły poziom powyżej 93%

Finalnie, model z dwiema warstwami ukrytymi o liczbie neuronów odpowiednio $N = 128$ oraz $N = 64$ został ostatecznie uznany za najbardziej efektywny, uzyskując średni poziom SROCC, celności oraz miary F1 równy 94%.



Rysunek 2 Estymacja rozkładu prawdopodobieństwa dla przykładowych próbek przez trzy najskuteczniejsze modele

Dodatkowym, praktycznym aspektem wdrożeniowym niniejszej pracy jest zaprojektowanie narzędzia bazującego na opracowanym estymatorze jakości QoE z przeznaczeniem dla działów obsługi klienta. Narzędzie to, przedstawione na rysunku 3, umożliwia predykcyjną ocenę jakości doświad-



Rysunek 3 Narzędzie do oceny jakości QoE w usłudze telewizyjnej z zaimplementowanym estymatorem QoE bazującym na subiektywnej ocenie jakości

czenia (QoE) użytkowników korzystających z usługi telewizyjnej Hiway. Analiza poziomu QoE dla wybranych dekoderów pozwala na odwzorowanie percepcji jakości przez użytkownika podczas konkretnej audycji, kanału lub całej sesji oglądania, to jest od uruchomienia do wyłączenia urządzenia. W rezultacie zaproponowane rozwiązanie oferuje dwie kluczowe korzyści: po pierwsze, zapewnia lepsze zrozumienie subiektywnych odczuć klientów zgłaszających problemy, a po drugie, umożliwia proaktywne wykrywanie urządzeń charakteryzujących się obniżonym poziomem QoE, co z kolei pozwala na szybsze działania naprawcze i interwencje, zanim problem zostanie zgłoszony przez użytkownika usługi.

Publikacje autora

Artykuły w czasopismach naukowych

1. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Subiektywne postrzeganie pojedynczych zniekształceń w telewizji opartej o strumieniowanie adaptacyjne*, *Przegląd Telekomunikacyjny – Wiadomości Telekomunikacyjne*, 2024, nr 4, s. 249–252.

2. P. Andruloniw, K. Kowalik, B. Partyka, P. Zwierzykowski: *Multi-layer QoE Learning System Implemented by Fiberhost*, *Applied Sciences*, 2023, vol. 13, iss. 4, . 2300-1–2300-23.

3. K. Kowalik, P. Andruloniw, B. Partyka, P. Zwierzykowski: *Telecom Operator's Approach to QoE*, *Journal of Telecommunications and Information Technology*, 2022, nr 2, s. 26–34.

4. P. Andruloniw, K. Kowalik, P. Zwierzykowski: *Unsupervised Learning Data-Driven Continuous QoE Assessment in Adaptive Streaming-Based Television System*, *Applied Sciences*, 2022, vol. 12, iss. 16, s. 8288-1–8288-23.

Rozdziały w monografiach i publikacje konferencyjne

1. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Quality of Experience Studies in TV Systems: Insights and Lessons Learned*, w: *Artificial Intelligence and Machine Learning. 43rd IBIMA Conference, Madrid, 2024*, Springer Nature, 2025, s. 267–273.

2. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Subiektywna ocena jakości wideo w sieci dostawcy dla usługi telewizyjnej opartej o strumieniowanie adaptacyjne*, w: *Perspektywy rozwoju w naukach inżynieryjno-technicznych – trendy, innowacje i wyzwania. Tom 1*, Wyd. Naukowe TYGIEL, 2024, s. 104–114.

3. P. Andruloniw, P. Zwierzykowski, K. Kowalik: *Quality of Experience Data Gathering in OTT Television Service*, w: *Proceedings of the 41st IBIMA Conference, Granada, 2023*, IBIMA Publishing, 2023, s. 119–122.

4. P. Andruloniw, K. Kowalik, P. Zwierzykowski: *Analiza skupień dla skończonego oraz ciągłego strumienia danych*, w: *Rozwiązania technologiczne XXI wieku – skutki i perspektywy rozwoju. Tom 1*, Wyd. Naukowe TYGIEL, 2022, s. 7–22.

5. P. Andruloniw, P.Zwierzykowski, K. Kowalik: *Analysis of Adaptive Video Streaming Degradation Based on Clustering Approach*, w: *Proceedings of the 39th IBIMA Conference, Granada, 2022*, IBIMA Publishing, 2022, s. 192–199.

6. P. Andruloniw, P. Zwierzykowski: *W jaki sposób metody uczenia maszynowego mogą pomóc w wykrywaniu degradacji i szacowaniu jakości usług wideo w sieciach telekomunikacyjnych?*, w: *Technologie XXI wieku – aktualne problemy i nowe wyzwania. Tom 3*, Wyd. Naukowe TYGIEL, 2021, s. 132–141.

Suma punktów MEiN: 350, **Suma punktów CORE: 530**