

Inteligentny system wspomagający proces identyfikacji perspektywicznych horyzontów w wielohoryzontowych złożach gazu ziemnego uwzględniający kryterium ekonomiczne ich udostępnienia i eksploatacji

Intelligent system supporting the process of identification of perspective horizons in multi-horizontal gas deposits taking into account economic criteria, their completion and exploitation

Adam Pańko

ORLEN Upstream Sp. z o.o.

STRESZCZENIE: W artykule zaprezentowano inteligentny system wspomagający proces identyfikacji perspektywicznych horyzontów złożowych w wielohoryzontowych złożach gazu ziemnego, uwzględniający kryterium ekonomiczne ich udostępnienia i eksploatacji. W procesie projektowania systemu zostały wykorzystane dotychczasowe doświadczenia firmy ORLEN Upstream z prac prowadzonych na obszarze zapadliska przedkarpackiego w utworach miocenu, obejmujące etap poszukiwania i eksploatacji wielohoryzontowych złóż gazu ziemnego. System został opracowany na bazie sztucznej inteligencji (SI) z wykorzystaniem między innymi sztucznych sieci neuronowych (SSN) i metod uczenia maszynowego (ML) oraz dodatkowo metod tzw. eksperymentu projektowanego (ang. *design of experiment*, DOE). Pierwsza część systemu obejmuje procesy związane z selekcją odpowiednich danych wejściowych i ich przygotowaniem do wykorzystania w kolejnych elementach systemu. Kolejnym etapem inteligentnego systemu jest identyfikacja perspektywicznych horyzontów złożowych w nowo wierconych odwiertach na podstawie wyników wykonanych opróbowań typu DST (ang. *drill stem test*) i testów produkcyjnych w dotychczas odwierconych i eksploatowanych odwiertach przez ORLEN Upstream. Następnym elementem systemu stanowi projekt bazy danych wejściowych do budowy zastępczego modelu złożowego (ZMZ). Do konstrukcji bazy danych wykorzystano metodę *Latin hypercube* i symulator numeryczny Eclipse. W dalszej części systemu skonstruowany model zastępczy został użyty do probabilistycznego generowania profili wydobywania gazu ze zidentyfikowanych w poprzednim etapie perspektywicznych horyzontów złożowych. Ostatnim elementem zaprojektowanego systemu jest analiza ekonomiczna opłacalności procesu udostępniania i eksploatacji, bazująca między innymi na wyznaczonych profilach wydobywania gazu. Wynikiem analizy jest wyznaczenie podstawowych wskaźników ekonomicznych inwestycji. Na podstawie przeprowadzonej analizy ekonomicznej tworzony jest ranking zidentyfikowanych horyzontów i podejmowana jest decyzja o ewentualnym udostępnieniu i eksploatacji zidentyfikowanego horyzontu lub odstąpieniu od jego opróbowania.

Słowa kluczowe: sztuczna inteligencja, uczenie maszynowe, sztuczne sieci neuronowe, zastępczy model złożowy, analiza ekonomiczna.

ABSTRACT: The article presents an intelligent system supporting the process of identification of perspective horizons in multi-horizontal gas deposits taking into account economic criteria of their completion and exploitation. Artificial Intelligence has been used for more than two decades as a development tool for solutions in several areas of the E&P industry: production control and optimization, forecasting, and simulation, among many others. The intelligent system was designed based on so far carried out work by the ORLEN Upstream company in the area of the Carpathian Foredeep (Miocene formations), including the phase of exploration and exploitation of multi-horizontal gas deposits. The system was developed based on artificial intelligence (AI) using, among other things, artificial neural networks (ANN), machine learning (ML), and additional methods of design of experiment (DOE). The first part of the designed system includes processes connected with the selection of proper input data and their preparation to be utilized in the next section of the system. The next stage of the intelligent system is the identification of perspective horizons in the new drilling wells based on results from performed DST and production tests in so far drilled and exploited wells by ORLEN Upstream. The subsequent stage is the design of input database for the construction of the Surrogate Reservoir Model (SRM). This input database was prepared using the Latin Hypercube method and the commercial reservoir simulator Eclipse. In the duration of the next stage of the system, the previously

Autor do korespondencji: A. Pańko, e-mail: adam.panko@orlen.pl

Artykuł nadesłano do Redakcji: 06.09.2022 r. Zatwierdzono do druku: 08.11.2022 r.

prepared Surrogate Reservoir Model was utilized to probabilistically generate production gas profiles from identified horizons. The final part of the intelligent system is the economic profitability analysis of investments, among other things, based on generated production profiles. The results of the economic analysis are economic indicators of investment. The decision concerning the possible completion and exploitation of the identified horizon or renouncement of the execution of the drill stem test is made on the basis of the economic results.

Key words: artificial intelligence, machine learning, artificial neural network, surrogate reservoir model, economic analysis.

Wprowadzenie

Rosnące zapotrzebowanie na gaz ziemny niesie ze sobą coraz większy nacisk na rozwój nowych, bardziej skutecznych metod poszukiwania, udostępniania i eksploatacji złóż tego surowca. Kolejnym czynnikiem znacząco wpływającym na konieczność rozwoju nowych metod interpretacyjnych są szybko wzrastające koszty związane z serwisami wykorzystywanymi w procesie poszukiwania złóż węglowodorów, w szczególności: usługi wiertnicze, geofizyczne, usługi związane z oprobowaniem i testami produkcyjnymi oraz zagospodarowaniem złóż. Szczególną rolę zaczyna odgrywać czas trwania poszczególnych operacji związanych z wykonaniem pomiarów geofizycznych i ich interpretacją oraz wykonaniem oprobowania i testów produkcyjnych perspektywicznych interwałów. Powyższe elementy nabierają szczególnego znaczenia w przypadku złóż wielohoryzontowych, dla których niejednokrotnie zachodzi potrzeba wykonania oprobowania kilku lub nawet kilkunastu potencjalnych horyzontów złożowych, co z kolei związane jest ze zwiększonymi kosztami wiercenia i udostępnienia do eksploatacji. Warto zaznaczyć, że z tych kilku/kilkunastu testowanych horyzontów, nawet w przypadku ich pozytywnej weryfikacji, tylko dwa lub trzy ze względu na możliwości jednoczesnego udostępnienia mogą być przeznaczone do eksploatacji, a pozostałe ewentualnie mogą być eksploatowane po przeprowadzeniu rekonstrukcji odwiertu i ich udostępnieniu w wyniku wykonania perforacji rur okładzinowych. W związku z tym zachodzi konieczność nie tylko zidentyfikowania horyzontów zapewniających uzyskanie przemysłowego przyływu gazu ziemnego, lecz także wykonania ich klasyfikacji, biorąc pod uwagę aspekty finansowe, czyli uzyskanie jak najszybszego zwrotu poniesionych nakładów.

Zaproponowanym w niniejszym artykule rozwiązaniem powyższego problemu jest wykorzystanie sztucznej inteligencji, która w ostatnich kilkunastu latach została z powodzeniem zaimplementowana do rozwiązywania różnych, nieraz bardzo złożonych zagadnień w wielu gałęziach gospodarki, także w branży naftowej E&P (Tadeusiewicz, 1993). Bardzo szerokie zastosowanie znalazły sztuczne sieci neuronowe (SSN) oraz uczenie maszynowe (ang. *machine learning*), wykorzystywane zarówno w zagadnieniach klasyfikacyjnych, jak i regresyjnych. W światowej oraz krajowej literaturze można znaleźć

różnorodne przykłady zastosowań sztucznej inteligencji, między innymi w procesie interpretacji profilowań geofizycznych, identyfikacji elektrofacji, do określenia nasycenia węglowodorami, podczas optymalizacji wydobywania i magazynowania węglowodorów (Pańko, 2008; Łętkowski et al., 2010) oraz jako główny element zastępczy modeli złożowych (Ali, 1994; Mohaghegh, 1995, 2018; Mohaghegh et al., 2006; Fernandes, 2012; Franti i Sieranoja, 2019; Barbacki, 2021).

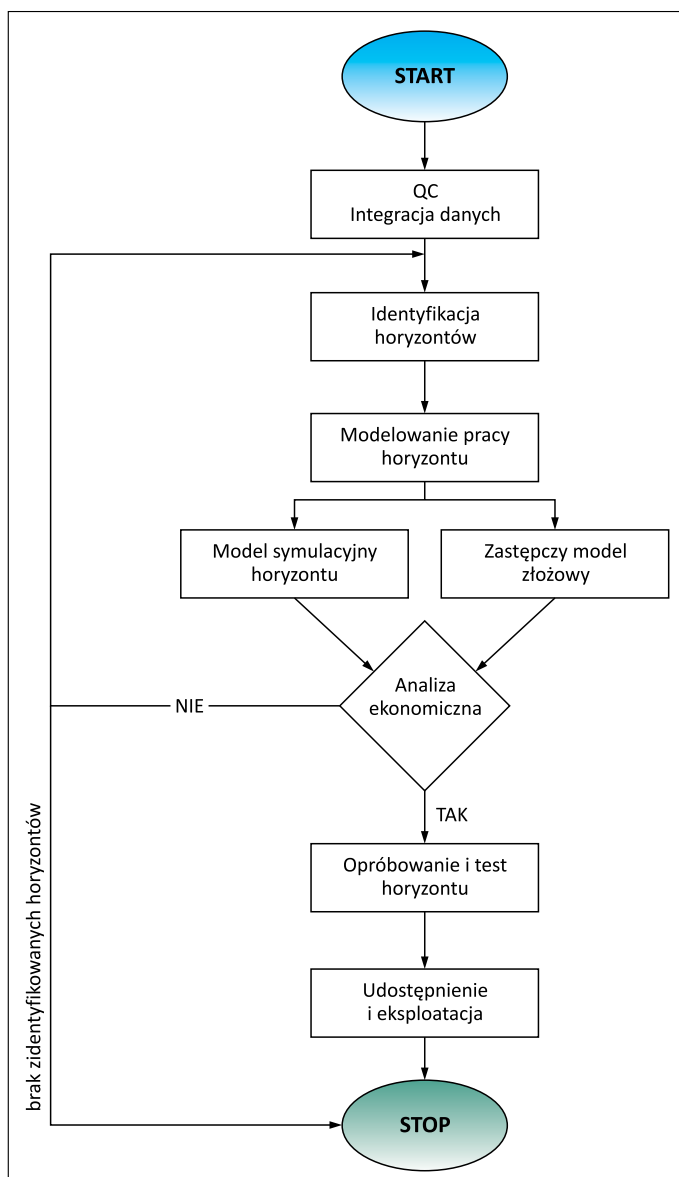
Do identyfikacji perspektywicznych interwałów złożowych w utworach miocenu zapadliska przedkarpackiego opracowano inteligentny system wykorzystujący między innymi metody nienadzorowane i nadzorowane uczenia maszynowego oraz sztuczne sieci neuronowe. Inteligentny system został zaprojektowany przy wykorzystaniu programów komputerowych MATLAB i Interactive Petrophysics oraz bibliotek obliczeniowych: Pandas, Sklearn i Matplotlib. Algorytm działania inteligentnego systemu przedstawiono na rysunku 1.

Algorytm działania inteligentnego systemu

Na rysunku 1 zaprezentowano poszczególne elementy inteligentnego systemu identyfikacji perspektywicznych interwałów złożowych w profilu odwiertu, wykorzystującego między innymi dane z pomiarów geofizycznych wykonanych w interwale niezarurowanym rurami okładzinowymi. Zaprojektowany system jest wieloelementowy, rozpoczyna się wyborem danych wejściowych, ich integracją i korelacją poprzedzoną procesem QC, a kończy – decyzją o udostępnieniu zidentyfikowanego horyzontu do eksploatacji.

Wybór danych wejściowych

Pierwszym bardzo ważnym etapem analizy jest wybór danych wejściowych. Głównym kryterium wyboru danych jest ich reprezentatywność, wiarygodność i odpowiednia ilość dla zastosowania różnorodnych algorytmów sztucznej inteligencji – między innymi *machine learning* (ML) i sztucznych sieci neuronowych (SSN). Każdy typ wykorzystanych danych powinien mieć istotny wpływ na proces identyfikacji perspektywicznych horyzontów złożowych (Mohaghegh, 2018). Wykorzystano następujące dane pochodzące z czterech odwiertów (W1,



Rysunek 1. Algorytm działania inteligentnego systemu

Figure 1. Intelligent system algorithm

W2, W3, W4) odwierconych w utworach miocenu zapadliska przedkarpackiego:

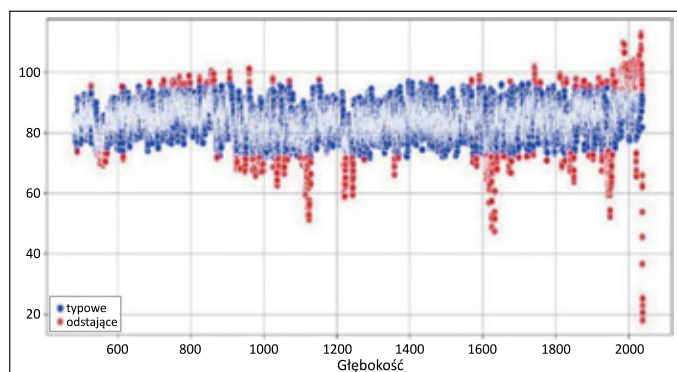
- wyniki pomiarów geofizyki odwiertowej:
 - GR – profilowanie naturalnej promieniotwórczości gamma [API],
 - NPHI – porowatość neutronowa [%],
 - RHOB – gęstość objętościowa [g/cm^3],
 - RT – oporność całkowita [om],
 - PE – indeks absorpcji fotoelektrycznej [b/e],
 - DT – profilowanie akustyczne czasu interwałowego [us/ft],
 - POTA – profilowanie potasu [%],
 - URAN – profilowanie uranu [%],
 - THOR – profilowanie toru [%];
- wyniki analiz laboratoryjnych rdzeni wiertniczych i próbek okruchowych oraz wskazania profilowania gazowego;

- wyniki analiz laboratoryjnych pobranych próbek płynów złożowych;
- wyniki wykonanych DST i testów produkcyjnych.

Powyższy zestaw danych został poddany odpowiedniej obróbce, tzw. preprocessingowi danych, czyli przygotowano je do wykorzystania przez sieci neuronowe i metody uczenia maszynowego. Przygotowanie danych wejściowych obejmowało następujące elementy:

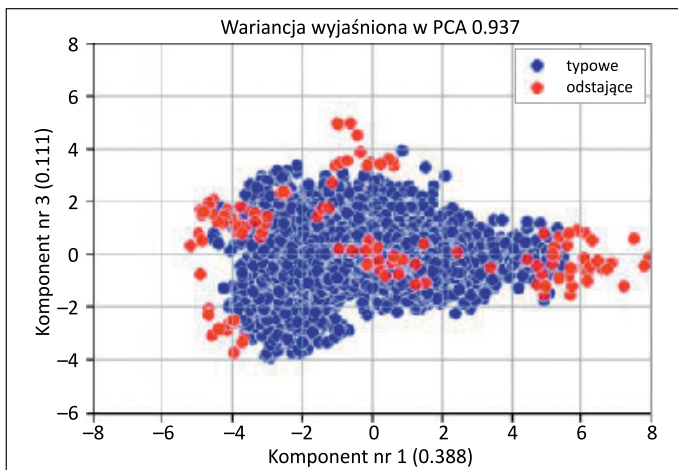
1. **Oczyszczenie** – a) usunięcie danych, które nie są istotne dla procesu identyfikacji perspektywicznych horyzontów (np. leżące poza zakresem głębokościowym, dla którego identyfikowane są perspektywiczne interwały; obejmujące wyniki opróbowań nieudanych pod względem technicznym, tzn. dla których nie uzyskano wystarczającej szczelności zapięcia pakera izolującego lub wystąpiła nieszczelność wykorzystywanego przewodu wiertniczego, itp.) oraz nie będą wykorzystane w procesie uczenia maszynowego, a także w dalszych analizach; b) usunięcie wierszy zawierających wartości tzw. *null* lub niezawierających żadnych wartości.

2. **Detekcja i usunięcie pomiarów odstających „outliers”** – bardzo istotny element algorytmu, pozwalający na identyfikację i usunięcie danych odstających, które w znaczący sposób wpływają na zniekształcenie danych wejściowych i są bardzo dużym utrudnieniem w procesie uczenia sieci neuronowych czy wykorzystaniu metod uczenia maszynowego. Etapem wstępnym jest określenie wpływu zmiany średnicy odwiertu wiertniczego na wyniki pomiarów geofizycznych na podstawie profilowania średnicy (*caliper log*). Drugi etap to wykorzystanie odpowiedniej metody detekcji pomiarów odstających. W celach porównawczych zastosowano trzy algorytmy: EllipticEnvelope (rysunek 2), Grubbsa i OneClassSVM. Pierwsze dwa algorytmy w uproszczeniu polegają na sprawdzeniu odległości każdego pomiaru od ogólnej średniej uwzględniającej wszystkie zmienne ze zbioru danych. Trzecia metoda, OneClassSVM (Boschetti i Massaron, 2017) (rysunek 3),



Rysunek 2. Wyniki działania algorytmu EllipticEnvelope

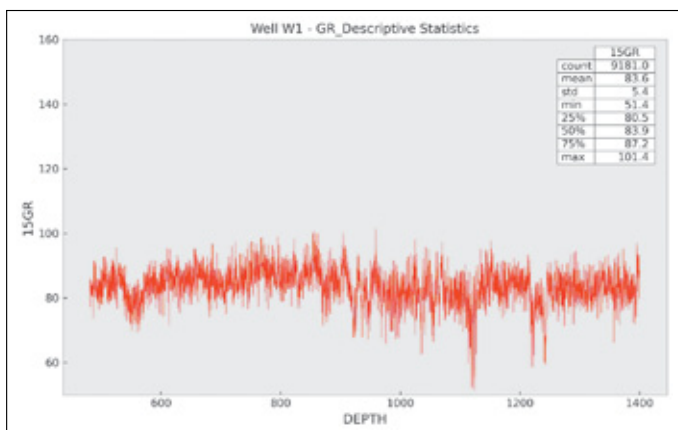
Figure 2. Result of EllipticEnvelope algorithm implementation



Rysunek 3. Wyniki działania algorytmu OneClassSVM
 Figure 3. Result of OneClassSVM algorithm implementation

to algorytm z obszaru uczenia maszynowego określający rozkład cech na podstawie samych danych. Ten algorytm ma szerokie zastosowanie i można go stosować w rozmaitych sytuacjach, w których celem jest poszukiwanie danych odstających lub znalezienie nietypowych danych.

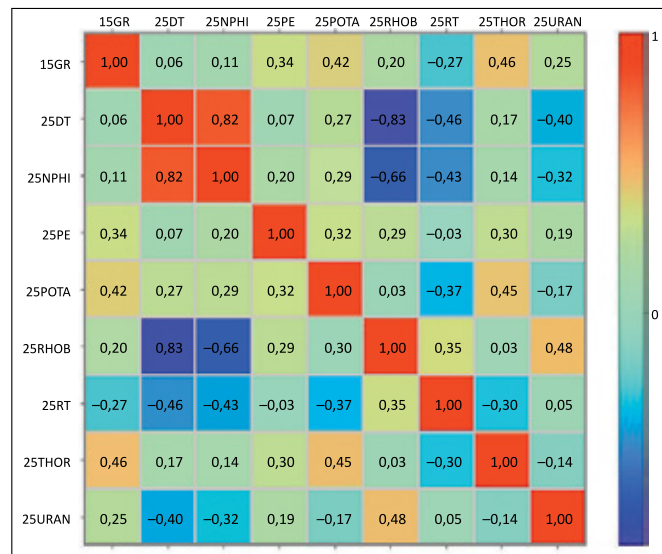
3. **Wizualizacja i statystyka opisowa** (ang. *descriptive statistics*) – wyznaczenie podstawowych wielkości statystycznych zbioru danych, takich jak: liczba pomiarów, wartość średnia, odchylenie standardowe, wartość minimalna, maksymalna, skośność, percentyle (rysunek 4).



Rysunek 4. Wizualizacja pomiarów geofizycznych – statystyka opisowa dla odwiertu W1

Figure 4. Well log visualization – descriptive statistics for well W1

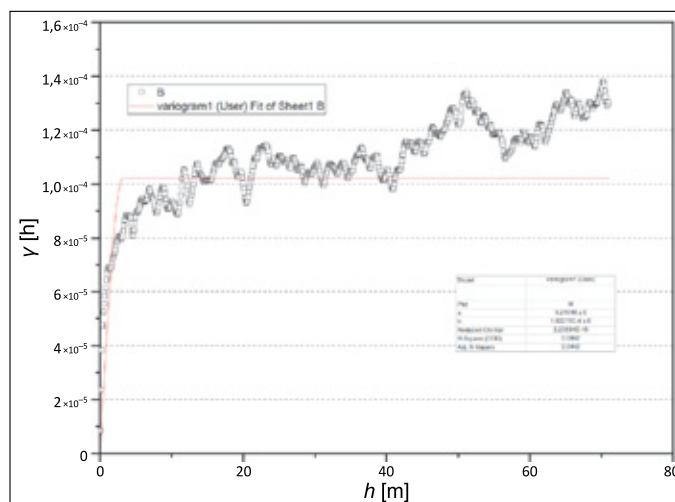
4. **Określenie korelacji pomiędzy poszczególnymi pomiarami** – wyznaczenie macierzy kowariancji i tzw. mapy korelacji (wsp. Pearsona) (rysunek 5). Stwierdzenie korelacji pomiędzy pomiarami jest bardzo istotnym elementem ze względu na fakt, że do analiz z wykorzystaniem uczenia maszynowego czy też sieci neuronowych korzystniej jest używać danych o słabszej korelacji.



Rysunek 5. Współczynnik korelacji pomiędzy pomiarami geofizycznymi – odwiert W1

Figure 5. Well logs correlation coefficient – well W1

5. **Analiza geostatystyczna** – wyznaczenie zasięgu korelacji danych pomiarowych (wariogramu) dla każdego zestawu danych pomiarowych. Pierwszym etapem jest konstrukcja wariogramu eksperymentalnego, do którego w drugim etapie dopasowywany jest model teoretyczny, na podstawie którego określany jest zasięg korelacji oraz wartość wariancji (Kwilosz et al., 2003). Obliczony zasięg korelacji dostarcza między innymi informacji odnośnie do potencjalnej miąższości wyznaczanych perspektywicznych horyzontów oraz jednorodności wyznaczonych interwałów (rysunek 6).
6. **Normalizacja danych pomiarowych** – ze względu na fakt, że dane pomiarowe mają różne skale/rozmiary, zostały one



Rysunek 6. Wariogram eksperymentalny z dopasowanym modelem teoretycznym – odwiert W1

Figure 6. Theoretical model matched to experimental variogram – well W1

sprowadzone do jednego wspólnego zakresu opisanego wartością średnią równą 0 i odchyleniem standardowym równym 1 (Fernandes, 2012) z wykorzystaniem wzoru (1):

$$x_n = \frac{x_i - \mu}{\sigma} \quad (1)$$

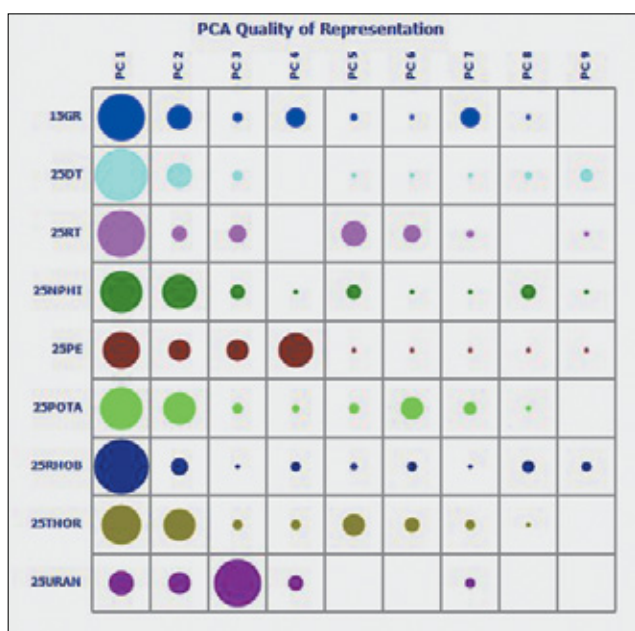
gdzie:

x_n – wartość znormalizowana,

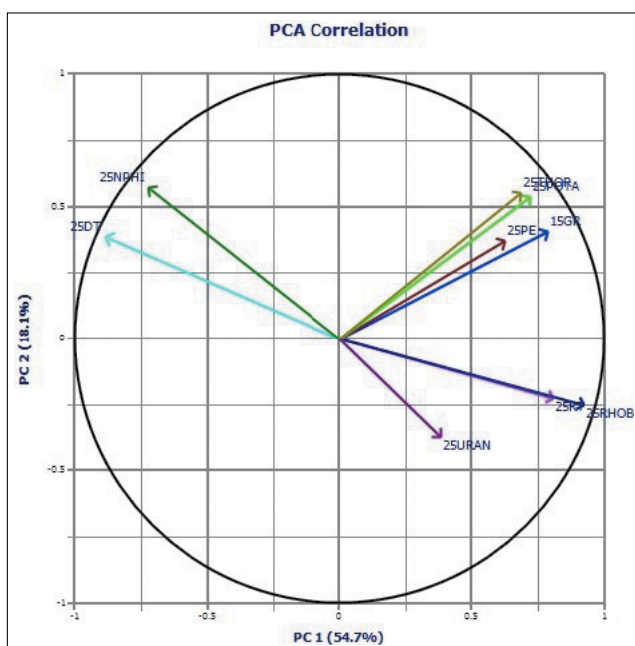
x_i – wartość pomiarowa,

μ – wartość średnia,

σ – odchylenie standardowe.



Rysunek 7. Analiza PCA – matryca reprezentatywności
Figure 7. PCA – quality of representation



Rysunek 8. Analiza PCA – korelacja
Figure 8. PCA – correlation

7. **Redukcja wielowymiarowości danych** – analiza czynników głównych (ang. *principal component analysis*, PCA), ograniczenie wielowymiarowości funkcji odpowiedzi przy jednoczesnym zachowaniu maksymalnej ilości informacji zawartej w oryginalnych danych pomiarowych. Ograniczenie wielowymiarowości jest szczególnie istotne dla etapu uczenia sztucznych sieci neuronowych czy metod uczenia maszynowego i powoduje znaczne skrócenie czasu trwania tego procesu. Bazując na wykonanej analizie PCA, wyznaczono, że 5 wektorów zawiera ponad 95% informacji ze wszystkich danych wejściowych, tzn. 95% wariacji wejściowej (rysunki 7 i 8).

Definicja perspektywicznych interwałów złożowych

Wykorzystując wyniki wykonanych próbowań typu DST (ang. *drill stem test*, opróbowanie przewierconych interwałów w trakcie wiercenia odwiertu) i testów produkcyjnych, we wszystkich wykonanych przez ORLEN Upstream odwiertach w utworach miocenu (z wyjątkiem odwiertu W1, który został wykorzystany do weryfikacji tworzonego systemu) wydzielono 6 typów interwałów, tzw. *flow units* (FU), z których wystąpił lub nie wystąpił dopływ medium złożowego do odwiertu. Wydzielone interwały zdefiniowane zostały według poniższego schematu:

- 1) FU_1 – interwał, w którym wykonano test produkcyjny i uzyskano przemysłowy przyływ gazu ziemnego;
- 2) FU_2 – interwał, w którym wykonano DST i uzyskano przyływ gazu ziemnego;
- 3) FU_3 – interwał, w którym wykonano DST i nie uzyskano przyływu medium złożowego;
- 4) FU_4 – interwał, w którym wykonano DST i uzyskano tylko przyływ wody złożowej;
- 5) FU_5 – interwał w stropie i spągu testowanego horyzontu definiowany jako uszczelniający;
- 6) FU_6 – pozostałe interwały niepoddane testowaniu, nieperspektywiczne.

Bazując na powyższej definicji, przypisano odpowiednie tzw. etykiety do danych uczących obejmujących pomiary geofizyczne (GR, NPHT, RHOB, RT, PE, DT, POTA, THOR, URAN) po procesie preprocessingu.

Identyfikacja perspektywicznych horyzontów przy wykorzystaniu sztucznych sieci neuronowych i metod uczenia maszynowego

Sztuczna inteligencja (SI) to zdolność maszyn do wykazywania ludzkich umiejętności, takich jak rozumowanie, uczenie

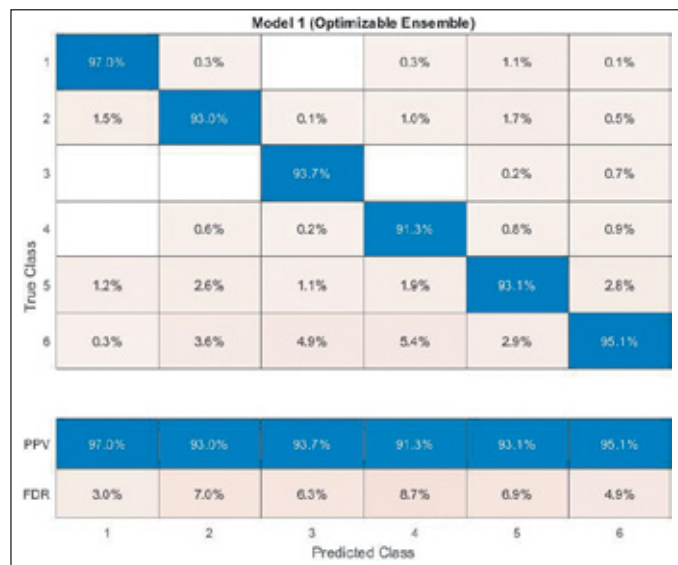
się, planowanie i kreatywność. Systemy SI są w stanie do pewnego stopnia dostosować swoje zachowanie, analizując skutki wcześniejszych działań i działając autonomicznie.

Uczenie maszynowe – samouczenie się maszyn albo systemy uczące się (ang. *machine learning*) to obszar sztucznej inteligencji poświęcony algorytmom, które poprawiają się automatycznie poprzez doświadczenie, czyli ekspozycję na dane. Algorytmy uczenia maszynowego budują model matematyczny na podstawie przykładowych danych, zwanych zbiorem uczącym, w celu prognozowania lub podejmowania decyzji bez bycia zaprogramowanym *explicite* przez człowieka do tego celu (Uczenie maszynowe).

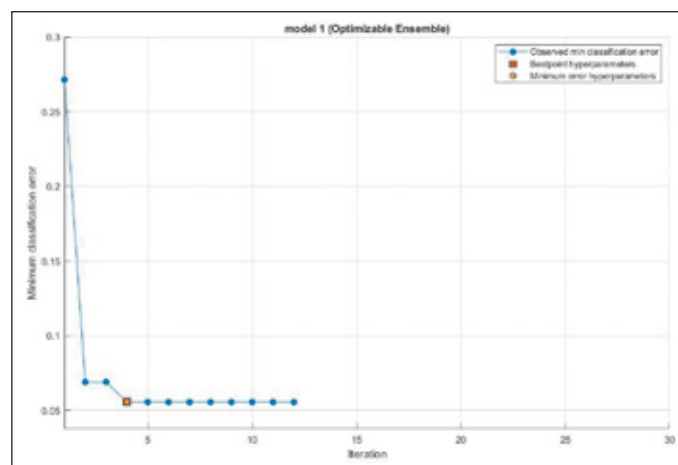
Bazując na przygotowanym (przy użyciu algorytmu do preprocessingu) zestawie danych uczących, wykorzystując różne metody uczenia maszynowego (ML) oraz sztuczne sieci neuronowe (SSN), wyznaczono perspektywiczne interwały złożowe. Zestaw danych uczących zawierał blisko 30 000 rekordów. Pierwszym etapem identyfikacji perspektywicznych interwałów było zastosowanie metod nadzorowanego uczenia maszynowego. W metodach tych razem ze zbiorem danych wejściowych dostarczana jest algorytmowi także oczekiwana odpowiedź (etykieta), tj. wyznaczone na podstawie uzyskanych wyników wykonanych DST i testów produkcyjnych interwały, tzw. *flow units* (FU). Dzięki takiemu nauczaniu oczekujemy, że po przedstawieniu nauczonemu modelowi nowych danych, które nie zostały wykorzystane do uczenia, algorytm prawidłowo zakwalifikuje dane (pomiar) do odpowiedniej kategorii (*flow units*).

Przetestowano 6 algorytmów uczących: Decision Trees (DT), Discriminant Analysis (DA), Naive Bayes Classifiers (NB), Nearest Neighbor Classifiers (KNN), Support Vector Machine (SVM), Ensemble Classifiers (EC) (Barbacki, 2021; MathWorks, 2021). Najlepsze wyniki otrzymano, stosując algorytm Ensemble Classifiers, uzyskując skuteczność klasyfikacji w porównaniu z wzorcem na poziomie około 94%. Algorytm ten testuje kilka modeli klasyfikatorów, a następnie tworzy nowy, który jest kombinacją testowanych modeli i wykazuje lepsze zdolności klasyfikacyjne niż każdy z osobna testowany klasyfikator. Na rysunku 9 zaprezentowano tzw. *confusion matrix*, przedstawiającą skuteczność klasyfikacji dla poszczególnych *flow units* (PPV – poprawnie sklasyfikowanych; FDR – błędnie sklasyfikowanych), a na rysunku 10 – wykres obrazujący minimalny błąd klasyfikacji vs. liczba iteracji.

W kolejnym etapie zastosowano sztuczne sieci neuronowe (SSN) do identyfikacji tzw. *flow units*. Wykorzystano ten sam zestaw danych wejściowych jak w przypadku metod uczenia maszynowego. Dane te zostały podzielone na trzy zestawy: uczący, walidacyjny i testujący, odpowiednio w procentowej proporcji 70 : 15 : 15. Zestaw uczący został wykorzystany do uczenia sieci neuronowej, zestaw walidacyjny – do tzw.



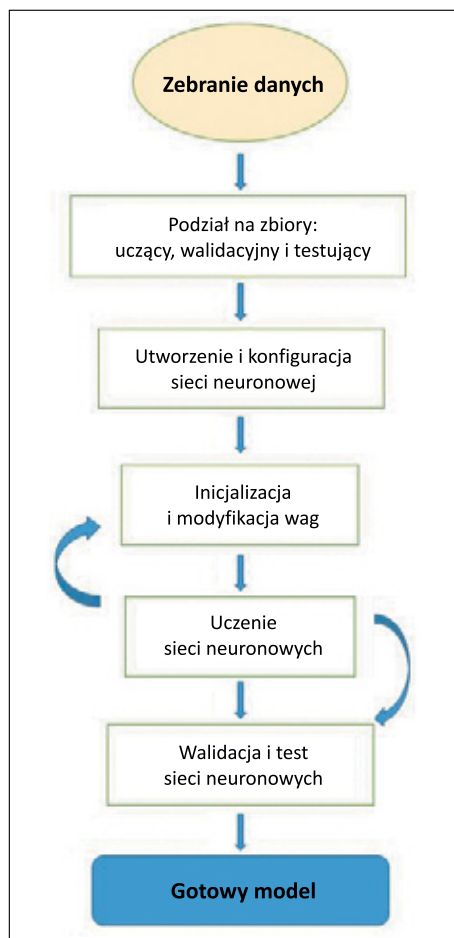
Rysunek 9. Confusion matrix – skuteczność metody
 Figure 9. Confusion Matrix – method effectiveness



Rysunek 10. Minimalny błąd klasyfikacji
 Figure 10. Minimum classification error

walidacji sieci i jego zadaniem było zapobieganie przeuczeniu się sieci, które powoduje w znaczącym stopniu redukcję zdolności generalizujących nauczonej sieci, natomiast zestaw testujący wykorzystano do potwierdzenia dopasowania nauczonej sieci oraz upewnienia się, czy proces walidacji przebiegł poprawnie, czy był tylko tzw. szczęśliwym trafem. Implementacja sieci neuronowych w procesie identyfikacji *flow units* przebiegała według algorytmu zaprezentowanego na rysunku 11.

Najlepsza skuteczność klasyfikacji/rozpoznawania została uzyskana dla sieci *four-layers feed-forward*, złożonej z warstwy wejściowej zawierającej 9 neuronów, z 4 warstw ukrytych zawierających odpowiednio 13, 15, 16 i 16 neuronów oraz warstwy wyjściowej zawierającej 6 neuronów odpowiadających odpowiedniej jednostce *flow units*. W procesie uczenia wykorzystano algorytm Levenberga–Marquardta, zapewniający stosunkowo dużą szybkość, a zarazem skuteczność uczenia.

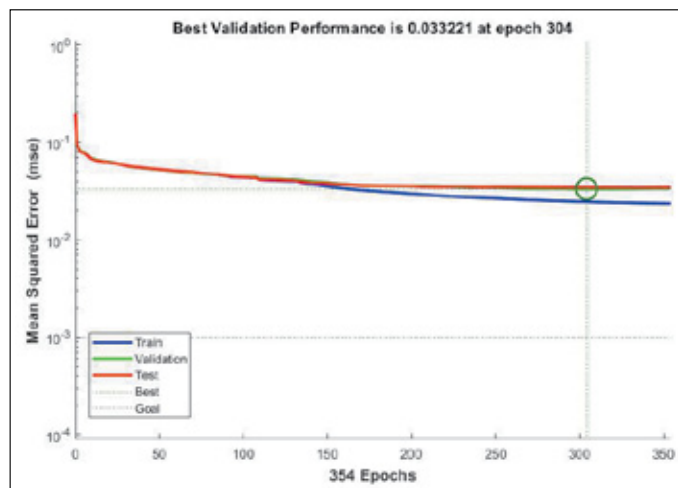


Rysunek 11. Algorytm analizy przy wykorzystaniu sztucznych sieci neuronowych

Figure 11. Algorithm of neural network implementation

Ostatecznie skuteczność rozpoznawania dla wszystkich zbiorów danych (uczący, walidacyjny i testujący) osiągnęła poziom prawie 90%, który można uznać za wysoki. Strukturę wykorzystanej sieci neuronowej zaprezentowano na rysunku 12, a na rysunku 13 – wykres przedstawiający proces uczenia się sieci neuronowych w postaci malejącego błędu uczenia (ang. *mean squared error*, MSE, czyli błąd średniokwadratowy wyrażony jako różnica kwadratów odchyleń).

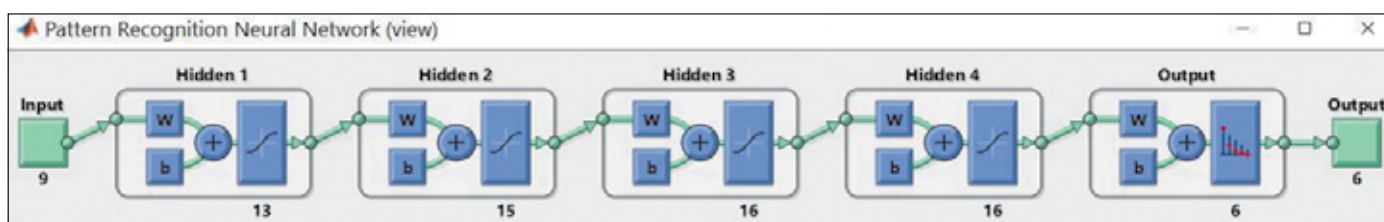
Wytrenowany model sieci neuronowych *four-layers feed-forward network* oraz model *machine learning – Ensemble Classifiers* zostały zastosowane do identyfikacji perspektywicznych horyzontów złożowych w odwiercie W1. Odwiert



Rysunek 13. Jakość uczenia – krzywe minimalizacji błędu uczenia dla zbioru uczącego, walidacyjnego i testowego

Figure 13. Quality of learning – minimization error curves for learning, validation and test sets

ten został wybrany do sprawdzenia działania opracowanego systemu, ponieważ dla tego odwiertu były dostępne wyniki przeprowadzonych próbowań typu DST i testów produkcyjnych oraz dodatkowo jest on wyposażony w selektywny zestaw wydobywczy i w sposób ciągły eksploatowany. W wyniku zastosowania zaprojektowanego systemu zidentyfikowano dziewięć perspektywicznych interwałów złożowych. Zidentyfikowane w wyniku działania systemu interwały bardzo dobrze korelowały z interwałami, w których zostało wykonane opróbowanie i testy produkcyjne w trakcie wiercenia odwiertu, co potwierdziło dużą skuteczność działania opracowanego systemu. Biorąc pod uwagę możliwości techniczne udostępnienia poszczególnych zidentyfikowanych horyzontów, zaprojektowano algorytm konsolidacyjny, którego zadaniem było pogrupowanie interwałów w większe skupiska, bazując na założonej tzw. bramce, czyli maksymalnej odległości pomiędzy interwałami, oraz wykorzystując wyznaczone na podstawie analiz laboratoryjnych rdzeni wiertniczych oraz krzywych geofizycznych parametry odcięcia, tzw. *cut-off*. Zadaniem algorytmu było także wyznaczenie dla skonsolidowanych horyzontów wartości minimalnej, maksymalnej oraz wartości średniej dla danej wielkości z profilowań geofizycznych. Wielkości te zostały wykorzystane w następnym etapie



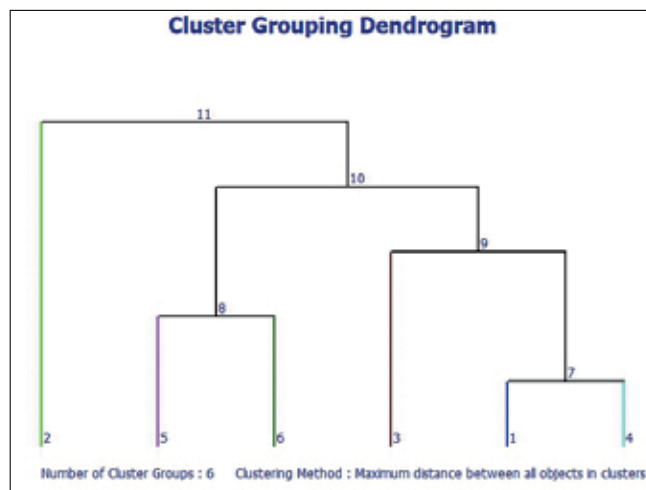
Rysunek 12. Architektura sztucznej sieci neuronowej

Figure 12. Neural networks architecture

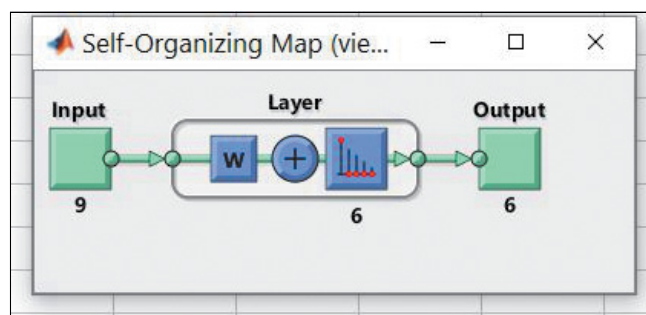
systemu do tworzenia zastępczego modelu złożowego (ZMZ) (Mohaghegh et al., 2006).

Dodatkowo w celu przeprowadzenia analizy porównawczej metod uczenia maszynowego nadzorowanego i nienadzorowanego w procesie identyfikacji tzw. *flow units* wykorzystano następujące metody uczenia nienadzorowanego: *K-means Clustering* oraz *Self-Organizing Map* (SOM) (Ali, 1994; Franti i Sieranoja, 2019). Zbiór danych wejściowych dla tych metod nie zawierał żadnych odpowiedzi (etykiet) i algorytm uczący sam klasyfikował dane pomiędzy poszczególne kategorie (*flow units*). Jediną dostarczoną algorytmowi informacją była wskazówka odnośnie do liczby kategorii, na które powinien zostać podzielony zestaw danych wejściowych, czyli tak samo jak w przypadku metod nadzorowanych – 6 kategorii. Na rysunku 14 zaprezentowano dendrogram dla metody uczenia nienadzorowanego *K-means Clustering*, a na rysunku 15 – architekturę sieci *Self-Organizing Map*.

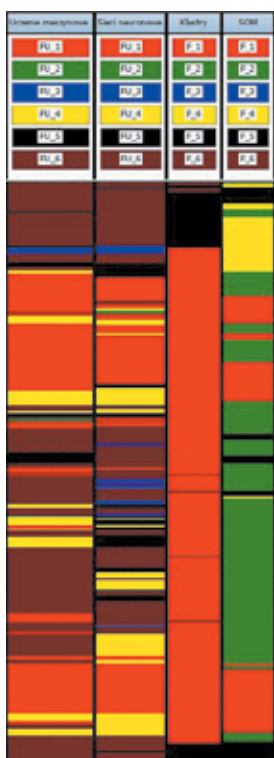
Porównanie uzyskanych wyników w postaci zidentyfikowanych perspektywicznych interwałów przy implementacji sztucznych sieci neuronowych oraz metod nadzorowanego i nienadzorowanego uczenia maszynowego przedstawiono na rysunkach 16 i 17 (oznaczenie poszczególnych ścieżek na rysunkach: Uczenie maszynowe – algorytm *Ensemble Classifiers*; Sieci neuronowe – sieć *four-layers feed-forward*; Klastry – *K-means Clustering*; SOM – *Self-Organizing Map*). Dodatkowo na rysunku 18 przedstawiono zidentyfikowany interwał zawodniony, który nie był przedmiotem dalszej analizy.



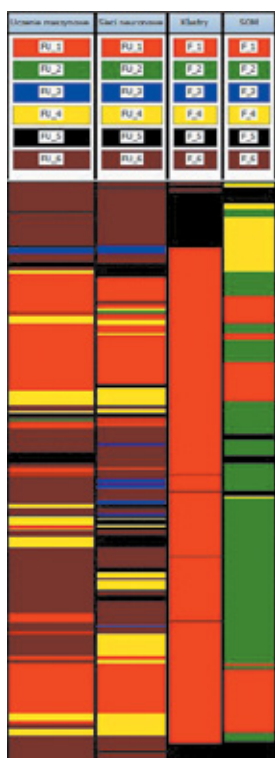
Rysunek 14. Dendrogram dla *K-means Clustering*
 Figure 14. K-means Clustering – dendrogram



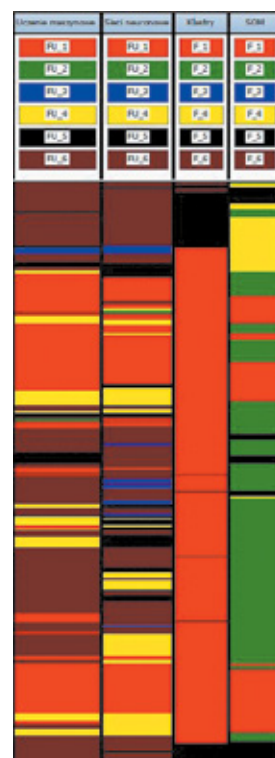
Rysunek 15. Architektura sieci SOM
 Figure 15. SOM network architecture



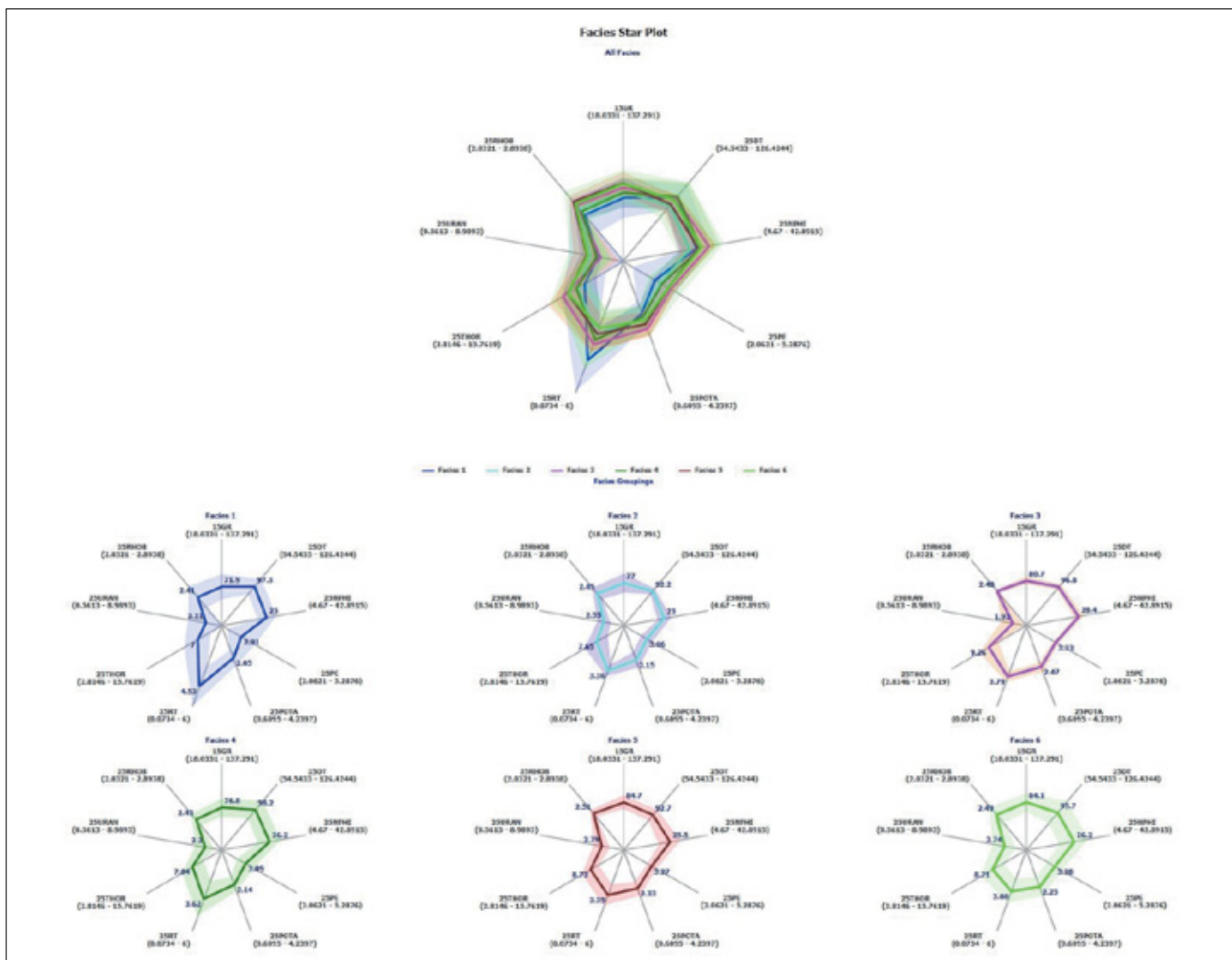
Rysunek 16. *Flow units* – interwał 1
 Figure 16. Flow units – interval 1



Rysunek 17. *Flow units* – interwał 2
 Figure 17. Flow units – interval 2



Rysunek 18. *Flow units* – zawodniony interwał
 Figure 18. Flow units – watered down interval



Rysunek 19. Wykres radarowy – wartości parametrów dla zidentyfikowanych perspektywicznych interwałów

Figure 19. Radar chart – parameter values for identified perspective intervals

Wielkości graniczne oraz średnie wartości poszczególnych parametrów geofizycznych dla danego interwału, tzw. *flow units*, zostały przedstawione w postaci tzw. wykresu radarowego na rysunku 19. Porównując perspektywiczne interwały zidentyfikowane na podstawie:

- sztucznych sieci neuronowych;
- modelu *Ensemble Classifiers*;
- *Self-Organizing Map* (SOM);
- *K-means Clustering*,

można zaobserwować pewne podobieństwo pomiędzy niektórymi interwałami zidentyfikowanymi za pomocą sztucznych sieci neuronowych, metod uczenia nadzorowanego i metody SOM, natomiast w przypadku większości interwałów widoczna jest rozbieżność z metodą *K-means Clustering*. Biorąc pod uwagę powyższe obserwacje, do dalszego wykorzystania w kolejnych etapach działania inteligentnego systemu zostały wybrane interwały zidentyfikowane za pomocą wytrenowanego modelu sztucznych sieci neuronowych. Interwały te najlepiej

odtwarzają horyzonty opróbowane/testowane w odwiercie W1, a zarazem potwierdzają bardzo dobrą zdolność predykcyjną wytrenowanego modelu. Wytrenowany model sieci neuronowych będzie także wykorzystywany do identyfikacji perspektywicznych interwałów w nowo wierconych odwiertach na podstawie pomiarów geofizyki wiertniczej.

Zastępczy model złożowy (ZMZ)

Kolejnym elementem zaprojektowanego inteligentnego systemu jest zastępczy model złożowy (ZMZ, w literaturze angielskiej *Surrogate Reservoir Model*, SRM (Mohaghegh et al., 2006)). Zastępczy model złożowy jest to replika klasycznych modeli numerycznych, które naśladują czy też odtwarzają ich możliwości z dużą dokładnością i z uwzględnieniem całej ich złożoności, do tego w nieporównywalnie krótszym czasie (Mohaghegh, 1995).

Zastępczy model złożowy powstaje na bazie konwencjonalnego numerycznego modelu złoża obejmującego dwie części: pierwsza to model geologiczny złoża, natomiast druga to tzw. model dynamiczny charakteryzujący przepływ płynów w złożu (Nagy et al., 2019).

Zastępczy model złożowy został wykorzystany do generowania profili wydobywania gazu ze zidentyfikowanych w poprzednim etapie perspektywicznych horyzontów. Profile te stanowią dane wejściowe do analizy ekonomicznej, której wyniki zostały w dalszym etapie wykorzystane do wyboru najlepszych horyzontów pod względem prognozowanego zwrotu inwestycji. Na podstawie uzyskanych wyników analizy ekonomicznej w wytypowanych horyzontach zostanie wykonane opróbowanie oraz test produkcyjny i zostaną one udostępnione do regularnego wydobywania gazu.

Pierwszym etapem budowy zastępczego modelu złożowego było określenie parametrów modelu, które mają istotny wpływ na wielkość produkcji gazu z modelowanego horyzontu. Wybór parametrów ma także istotne znaczenie dla tworzenia sieci neuronowej, która będzie trenowana na otrzymanych wynikach z modelu symulacyjnego. W celu uzyskania jak najlepszego wytrenowania sztucznej sieci neuronowej oraz zachowania jej zdolności generalizujących zdobytą wiedzę konieczne jest ograniczenie wymiarowości analizowanego zagadnienia, czyli redukcja parametrów, na podstawie których konstruowany będzie model symulacyjny i następnie zastępczy model złożowy. W celu określenia zestawu parametrów, które znacząco wpływają na wielkość produkcji gazu, wykonana została tzw. analiza wrażliwości, polegająca na modyfikacji jednego parametru w określonym zakresie przy niezmiennych pozostałych parametrach modelu bazowego i przeprowadzeniu analizy jego wpływu na wielkość wydobywania gazu. Bazując na analizach profili sumarycznego wydobywania, otrzymano zestaw 10 parametrów: porowatość, miąższość, nasycenie wodą, przepuszczalność, stosunek miąższości całkowitej do efektywnej (NTG), gradient ciśnienia złożowego, głębokość stropu interwału, wielkość strefy drenażu (długość i szerokość), początkowa wydajność odbioru gazu – który to zestaw został następnie wykorzystany do budowy zastępczego modelu złożowego.

Kolejnym etapem było wygenerowanie odpowiednio dużego zestawu danych wejściowych, które umożliwiają efektywne wytrenowanie sztucznej sieci neuronowej. Do tego celu wykorzystano metodę eksperymentu projektowanego *Latin hypercube* (McKay et al., 1979; Helton i Davis, 2002), umożliwiającą generowanie rozkładu parametrów wejściowych równomiernie pokrywających całą 10-parametrową przestrzeń. Wygenerowano 300 kompletów 10-parametrowych zestawów danych wejściowych, na których przy wykorzystaniu symulatora numerycznego Eclipse zostało przeprowadzonych

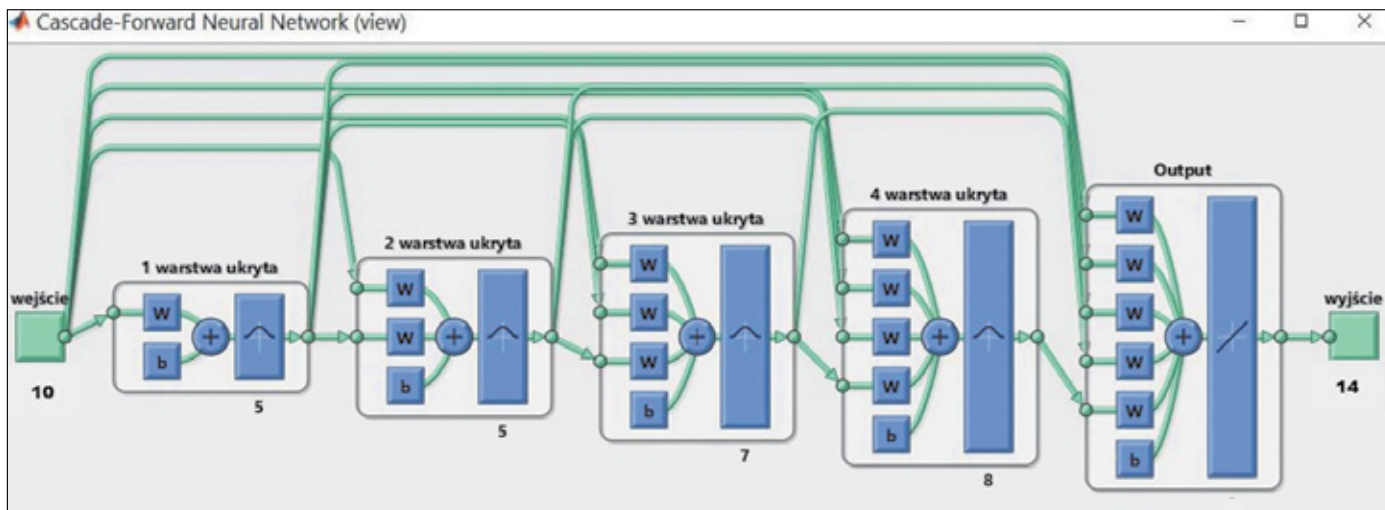
300 symulacji numerycznych. Uzyskane wyniki symulacji posłużyły jako zestawy danych do procesu nauki sztucznych sieci neuronowych stanowiących zastępczy model złożowy. Proces uczenia sieci neuronowych składał się z następujących elementów (Nagy et al., 2019):

- przyjęcie wstępnej konstrukcji sztucznej sieci neuronowej (SNN);
- uczenie sztucznej sieci neuronowej (SSN);
- ocena wyników generowanych przez SSN;
- modyfikacja struktury SSN;
- opracowanie nowych ciągów uczących w przypadku braku akceptowalności otrzymanych rezultatów.

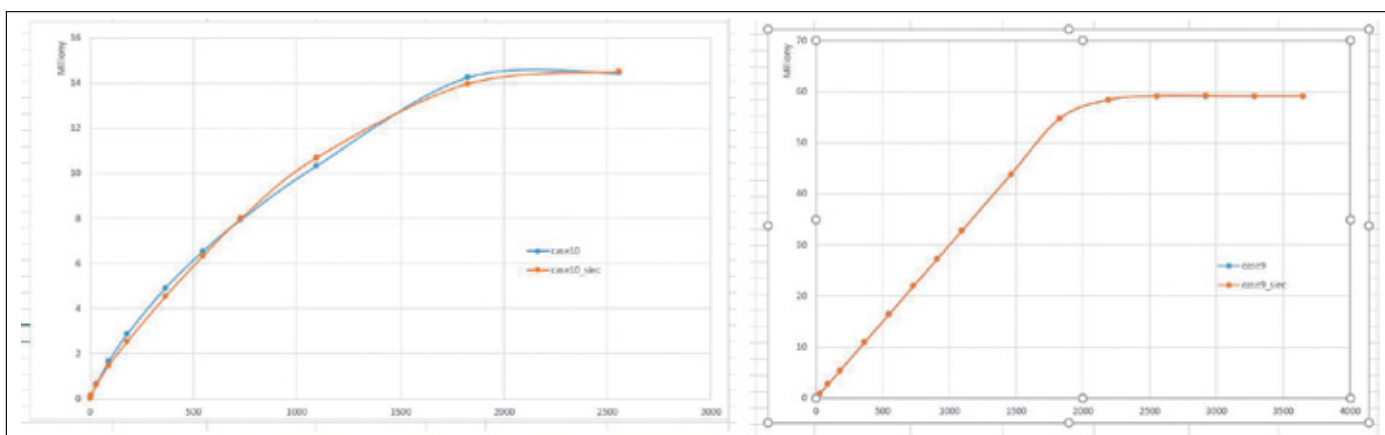
Dodatkowo po wytrenowaniu sieci neuronowej wykonano test na nowym zestawie danych, tj. 150 kompletach rozkładów danych wejściowych wygenerowanych metodą *Latin hypercube* (McKay et al., 1979; Helton i Davis, 2002), które nie zostały wcześniej wykorzystane w procesie uczenia sztucznej sieci neuronowej. Otrzymano średni błąd (różnica pomiędzy wynikami symulatora a odpowiedzią SNN) nieznacznie przekraczający 8%. Najlepsze wyniki zostały uzyskane dla sieci kaskadowej jednokierunkowej składającej się z 4 warstw ukrytych o liczbie neuronów odpowiednio 5, 5, 7 i 8 oraz warstwy wejściowej o 10 neuronach i warstwy wyjściowej o 14 neuronach obejmującej całą prognozowaną długość okresu wydobywania gazu, tj. 7 lat. Architekturę sztucznej sieci neuronowej przedstawiono na rysunku 20, a uzyskane dopasowanie pomiędzy wynikami symulatora i odpowiedzią SSN dla dwóch przykładowych zestawów danych wejściowych przedstawiono na rysunku 21.

Analiza ekonomiczna sposobu udostępnienia zidentyfikowanych perspektywicznych horyzontów

Wykorzystując zidentyfikowane interwały złożowe oraz wyznaczone dla nich zakresy wartości obejmujące pierwszą grupę parametrów (wykorzystanych do budowy zastępczego modelu złożowego), tj.: porowatość, przepuszczalność, nasycenie wodą, NTG, głębokość stropu horyzontu, a także dodatkowo bazując na interpretacji opróbowań wykonanych w odwiertach W2, W3 i W4, wyznaczono zakresy (minimalną i maksymalną wartość) dla drugiej grupy parametrów, tj. rozmiarów strefy drenażu, gradientu ciśnienia złożowego i początkowej wydajności odbioru gazu. W celu wykorzystania skonstruowanego w poprzednim etapie zastępczego modelu złożowego do generowania profili wydobywania gazu ze zidentyfikowanych horyzontów – zaprojektowano aplikację komputerową zaprezentowaną na rysunku 22. Biorąc pod uwagę aktualne techniczne możliwości udostępnienia jednocześnie maksymalnie dwóch interwałów złożowych, wykorzystując wyznaczone rozkłady parametrów wejściowych oraz bazując



Rysunek 20. Architektura kaskadowej sieci neuronowej
Figure 20. Cascade-Forward Neural Network architecture



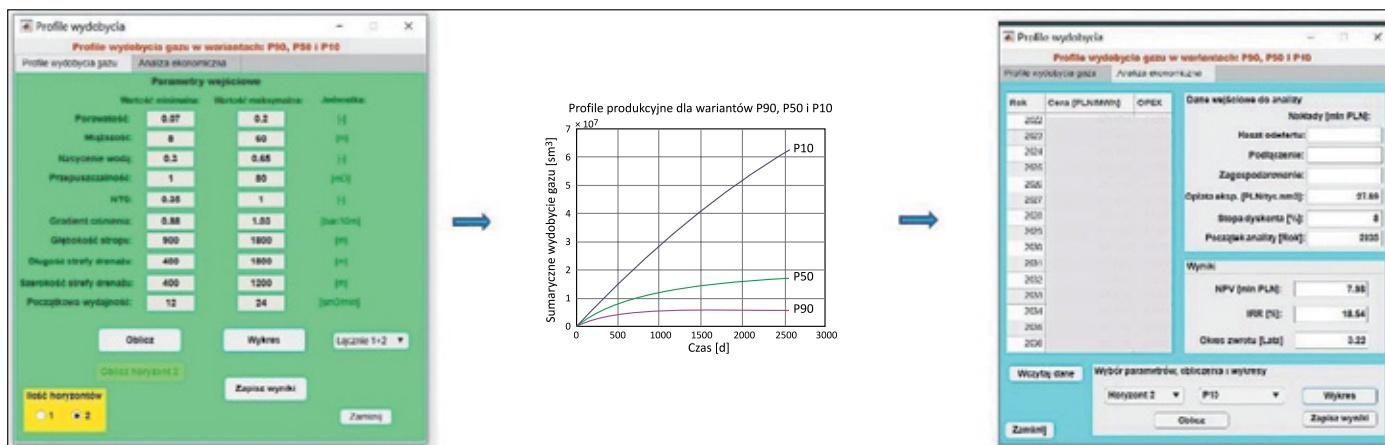
Rysunek 21. Porównanie odpowiedzi SSN z wynikami modelu symulacyjnego
Figure 21. Comparison of neural network results with simulation model outcome

na zastępczym modelu złożowym, przeprowadzono 2000 symulacji wydobywania gazu ziemnego z wyznaczonych interwałów. Uzyskane wyniki posłużyły do probabilistycznego wyznaczenia profili wydobywania gazu w wariantach P90, P50 i P10, które przedstawiono na rysunku 22 (Pańko, 2006). Skonstruowany program komputerowy pozwala wyznaczyć profile zarówno dla jednego, jak i dwóch równocześnie eksploatowanych horyzontów (eksploatowanych selektywnie).

Ostatnim elementem systemu jest wykonanie analizy ekonomicznej opłacalności procesu udostępnienia i eksploatacji zidentyfikowanych horyzontów – na podstawie otrzymanych profili wydobywania gazu przy implementacji zastępczego modelu złożowego oraz przy wykorzystaniu dodatkowych informacji, takich jak: koszt wiercenia odwiertu (w tym nakłady na wykonanie opróbowania i testu produkcyjnego wyznaczonych horyzontów), koszt zagospodarowania lub tylko koszt podłączenia do istniejącej instalacji, wysokość opłaty eksploatacyjnej, stopa dyskonta, koszty OPEX w ujęciu rocznym (uwzględniające

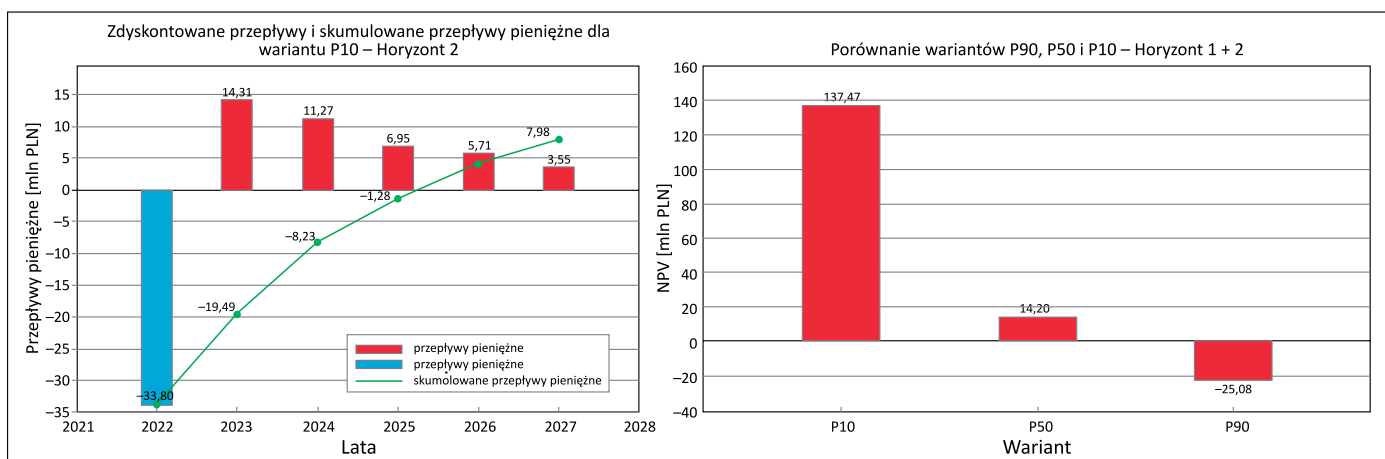
między innymi dzierżawę gruntu, podatki, utrzymanie załogi kopalni gazu ziemnego, wykorzystywane środki chemiczne do eksploatacji itp.), scenariusz cenowy sprzedaży gazu ziemnego w odpowiedniej perspektywie czasowej oraz data początku analizy, czyli rozpoczęcia regularnego wydobywania gazu. Bazując na powyższych parametrach, program wykonuje analizę ekonomiczną, wyznaczając następujące wskaźniki: wartość bieżącą netto (NPV), wewnętrzną stopę zwrotu (IRR) i okres zwrotu inwestycji (rysunek 22). Obliczane są także zdyskontowane przepływy i skumulowane przepływy pieniężne dla każdego analizowanego wariantu profili wydobywania gazu, tj. P90, P50 i P10 w scenariuszu dla jednego lub dwóch udostępnionych horyzontów złożowych (rysunek 23).

Na podstawie wyników przeprowadzonej analizy ekonomicznej opracowany został ranking zidentyfikowanych horyzontów, który potwierdził, że w pierwszej kolejności, biorąc pod uwagę prognozowany okres zwrotu inwestycji, powinny zostać udostępnione horyzonty, z których gaz ziemny



Rysunek 22. Aplikacja komputerowa – generowanie profili wydobycia gazu i analiza ekonomiczna

Figure 22. Computer application – generation of gas production profiles and economic analysis



Rysunek 23. Wyniki analizy ekonomicznej – zdyskontowane przepływy pieniężne i NPV dla wariantów zasobowych P90, P50 i P10

Figure 23. Economic analysis results – discounted cash flow and NPV for P90, P50 and P10 reserves cases

jest już eksploatowany w odwiercie W1. Fakt ten potwierdza poprawność działania zaprojektowanego inteligentnego systemu. Skonstruowany inteligentny system zostanie wykorzystany do identyfikacji perspektywicznych horyzontów w nowo wierconych odwiertach przez ORLEN Upstream oraz w procesie podejmowania decyzji odnośnie do kolejności ich udostępnienia do eksploatacji.

Podsumowanie i wnioski

Problematyka poruszona w artykule dotyczy możliwości wykorzystania sztucznej inteligencji w branży naftowej. Zastosowanie sztucznej inteligencji (SI) stanowi alternatywę dla dotychczas wykorzystywanych metod identyfikacji perspektywicznych horyzontów złożowych. W porównaniu z klasycznymi modelami symulacyjnymi warto podkreślić bardzo szybkie działanie wytrenowanego modelu (zastępuje model złożowy) na zebranych danych pochodzących

z nowo wierconych odwiertów. Z punktu widzenia ekonomii procesu poszukiwania i zagospodarowania złóż węglowodorów czas i skuteczność prowadzonych analiz mają coraz większe znaczenie ze względu na bardzo szybko wzrastające koszty serwisów wykorzystywanych podczas wykonywania tych prac.

Biorąc pod uwagę powyższe fakty, opracowany inteligentny system bardzo dobrze wpisuje się w bieżące trendy w branży naftowej. Pozwala on w szybki i stosunkowo prosty sposób zidentyfikować perspektywiczne horyzonty złożowe w profilu wierconego odwiertu oraz przy wykorzystaniu zastępczego modelu złożowego ocenić je pod względem możliwości zwrotu poniesionych nakładów na udostępnienie, zagospodarowanie i eksploatację. Daje możliwość wytypowania najlepszych pod względem ekonomicznym horyzontów złożowych w nowo wierconych odwiertach lub w już eksploatowanych, w których może zostać wykorzystany do identyfikacji perspektywicznych horyzontów i ich udostępnienia w trakcie wykonywania rekonstrukcji. Zaprojektowana aplikacja komputerowa pozwala

w bardzo intuicyjny sposób wykonać niezbędne analizy związane z: przygotowaniem danych wejściowych, wykorzystaniem sztucznych sieci neuronowych i metod uczenia maszynowego, budową zastępczego modelu złożowego oraz wykonaniem analizy ekonomicznej w celu podjęcia decyzji o udostępnieniu zidentyfikowanego horyzontu złożowego.

W celu zastosowania skonstruowanego inteligentnego systemu do innych obszarów, gdzie występują utwory miocenu, niezbędne jest zebranie nowych danych wejściowych oraz budowa nowego zastępczego modelu złożowego. Po wykonaniu tych niezbędnych modyfikacji system może być wykorzystany na nowych wielohoryzontowych złożach gazu ziemnego – już eksploatowanych lub będących w fazie rozpoznawania.

Opracowany system został przetestowany na eksploatowanym przez ORLEN Upstream wielohoryzontowym złożu gazu ziemnego na jednym z regularnie eksploatowanych odwiertów, potwierdzając skuteczność swojego działania.

Zestawienie wykorzystanych w artykule oznaczeń

SI/AI	– sztuczna inteligencja,
SSN	– sztuczne sieci neuronowe,
ML	– metody uczenia maszynowego,
DOE	– metoda eksperymentu projektowanego,
DST	– metoda opróbowania horyzontów w trakcie wiercenia odwiertu,
ZMZ/SRM	– zastępczy model złożowy,
NPV	– wartość bieżąca netto,
IRR	– wewnętrzna stopa zwrotu,
QC	– proces kontroli jakości danych,
PCA	– metoda czynników głównych,
FU	– opróbowany/przetestowany interwał w trakcie lub po zakończeniu wiercenia odwiertu,
NTG	– stosunek miąższości efektywnej do całkowitej,
Cut-off	– parametr odcięcia, np. dla porowatości,
P10, P50, P90	– klasyfikacja zasobów węglowodorów oszacowanych metodami probabilistycznymi, oznaczająca prawdopodobieństwo wystąpienia oszacowanej wielkości,
GR	– profilowanie naturalnej promieniotwórczości,
NPHI	– porowatość neutronowa,
RHOB	– gęstość objętościowa,
RT	– profilowanie oporności całkowitej,
PE	– indeks absorpcji fotoelektrycznej,
DT	– profilowanie akustyczne,
POTA	– profilowanie potasu,
URAN	– profilowanie uranu,
THOR	– profilowanie toru.

Literatura

Ali J.K., 1994. Neural Networks: A New Tool for the Petroleum Industry? *European Petroleum Computer Conference, Aberdeen, United Kingdom, March 1994*. DOI: 10.2118/27561-MS.

- Barbacki J., 2021. Zastosowanie sztucznej inteligencji w procesie klasyfikacji facji skał zbiornikowych w oparciu o dane geofizyki otworowej. *Wiadomości Naftowe i Gazownicze*, 2(267): 4–8.
- Boschetti A., Massaron L., 2017. Python. Podstawy nauki o danych. *Wydawnictwo Helion, Gliwice*.
- Fernandes M.A., 2012. Using Neural Networks for Determining Hydrocarbons Presence from Well Logs: A Case Study for Alagoas Basin. *Society of Petroleum Engineers*. DOI: 10.2118/153446-MS.
- Franti P., Sieranoja S., 2019. How much can k-means be improved by using better initialization and repeats? *Pattern Recognition*, 93: 95–112. DOI: 10.1016/j.patcog.2019.04.014.
- Helton J.C., Davis F.J., 2002. Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. *Scandia National Laboratories Albuquerque, New Mexico*. DOI: 10.2172/806696.
- Kwilosz T., Pańko A., Szott W., 2003. Zastosowanie metod geostatystycznych analizy danych dla potrzeb symulacji złożowych. *Prace Instytutu Górnictwa Naftowego i Gazownictwa*, 126: 1–35.
- Łętkowski P., Pańko A., Gołąbek A., 2010. Prognozowanie ciśnień dennych podziemnego magazynu gazu przy wykorzystaniu sztucznych sieci neuronowych. *Nafta-Gaz*, 66(7): 557–560.
- MathWorks, 2021. MATLAB documentation, R2021b.
- McKay M.D., Beckman R.J., Conover W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Computer Code. *American Statistical Association and the American Society for Quality Technometrics*, 42(1): 55–61.
- Mohaghegh S., 1995. Neural Network: What It Can Do for Petroleum Engineers. *Journal of Petroleum Technology*, 47(1): 42–42. DOI: 10.2118/29219-PA.
- Mohaghegh S., 2018. Data-Driven Analytics for the Geological Storage of CO₂. *Taylor & Francis Group*.
- Mohaghegh S., Modavi C., Hafez H.H., Haajizadeh M., Kenawy M., Guruswamy S., 2006. Development of Surrogate Reservoir Models (SRM) for Fast-Track Analysis of Complex Reservoirs. *Intelligent Energy Conference and Exhibition, Amsterdam, April 11–13*. DOI: 10.2523/99667-MS.
- Nagy S., Barbacki J., Dąbrowski K., Kaczmarek Ł., Klimkowski Ł., Kuczyński S., Łętkowski P., Maksimczuk M., Polański K., Siemek J., Skibiński J., Smulski R., Szott W., Wejrzanowski T., Włodek T., 2019. Zagadnienia eksploatacji niekonwencjonalnych złóż gazu ziemnego w skałach mułowcowo-lupkowych. *Wydawnictwa AGH, Kraków*.
- Pańko A., 2006. Probabilistyczne szacowanie zasobów złóż węglowodorów z wykorzystaniem metod geostatystycznych. *Nafta-Gaz*, 62(12): 655–668.
- Pańko A., 2008. Wykorzystanie możliwości sieci neuronowych w prognozowaniu i sterowaniu pracą podziemnego magazynu gazu (PMG). *Prace Instytutu Nafty i Gazu*, 147: 1–68.
- Tadeusiewicz R., 1993. Sieci neuronowe. *Akademicka Oficyna Wydawnicza RM, Warszawa*.
- Uczenie maszynowe. <https://pl.wikipedia.org/wiki/Uczenie_maszynowe> (dostęp: 20.06.2022).



Mgr inż. Adam PAŃKO
Koordynator Zespołu ds. Inżynierii Złożowej
ORLEN Upstream Sp. z o.o.
ul. Bielańska 12
00-085 Warszawa
E-mail: adam.panko@orlen.pl