Sprawozdanie 1

Rozdział 1 - Playground

Testowanie funkcji playground:

1a) Under-fitting

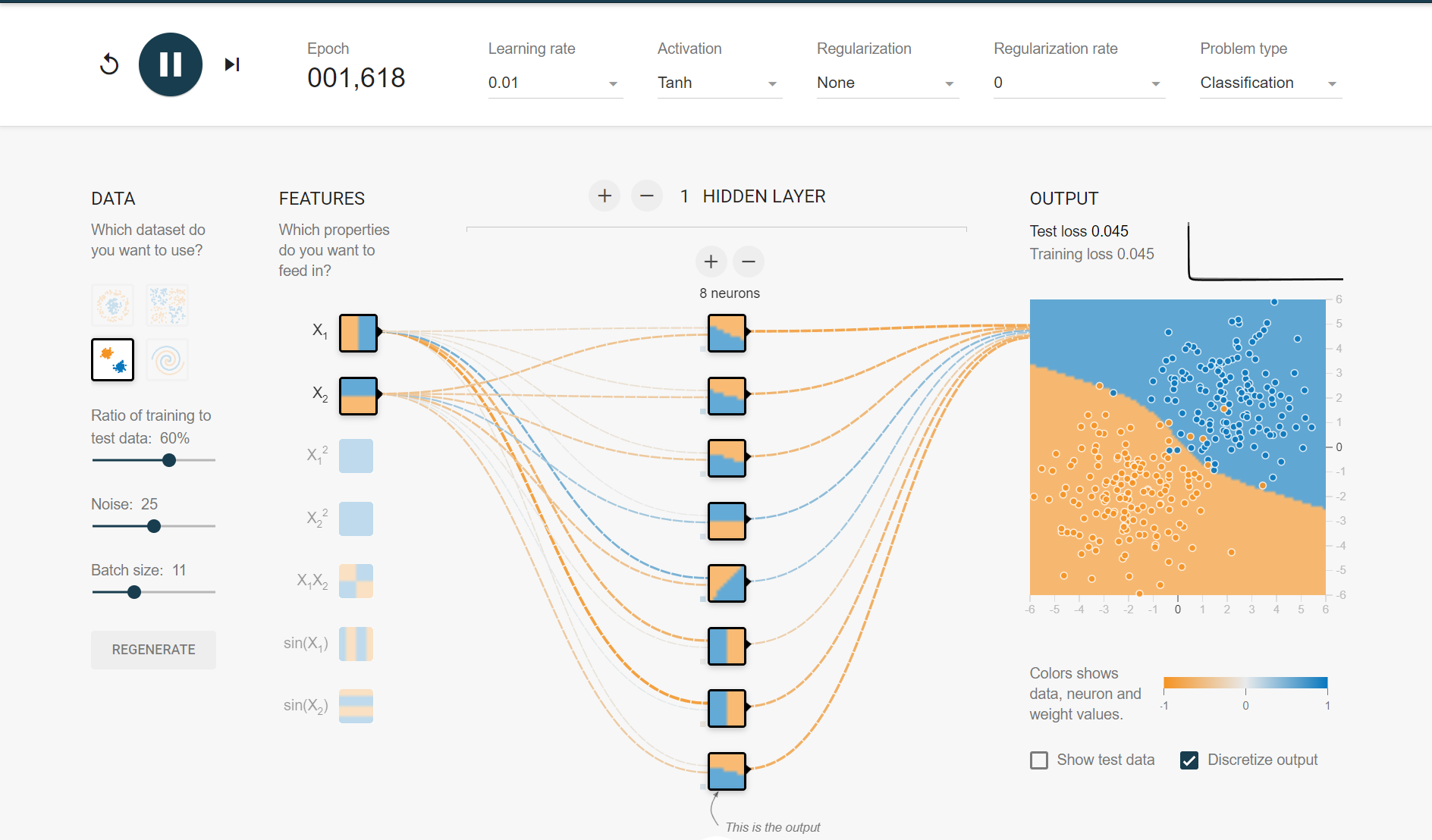
Obraz zawierający tekst, zrzut ekranu, diagram, oprogramowanie

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, diagram

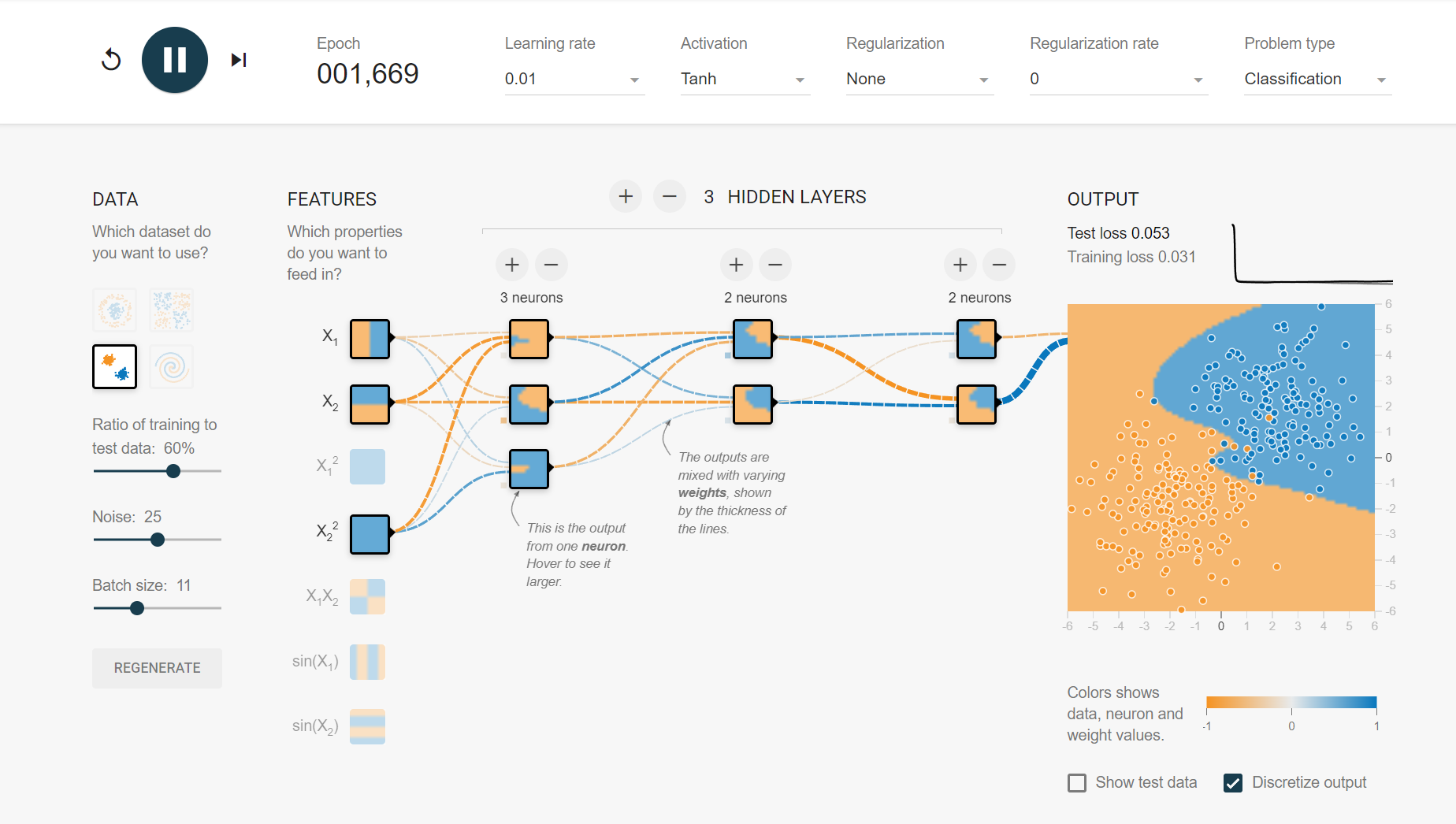
Opis wygenerowany automatycznie

1b) Over-fitting

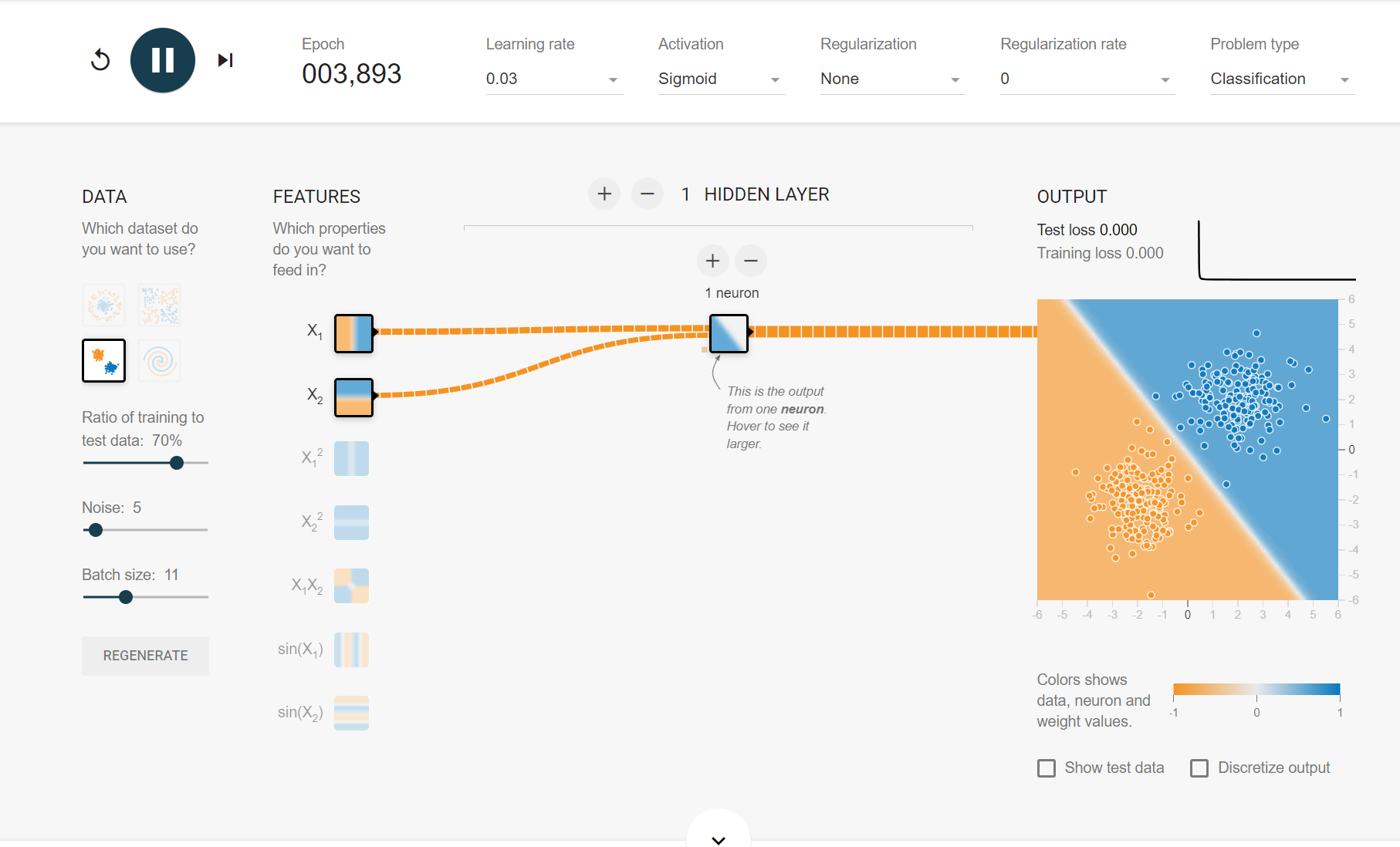


Obraz zawierający tekst, zrzut ekranu, diagram, mapa

Opis wygenerowany automatycznie



1c) Appropriate-fitting (Najlepsza minimalna architektura)

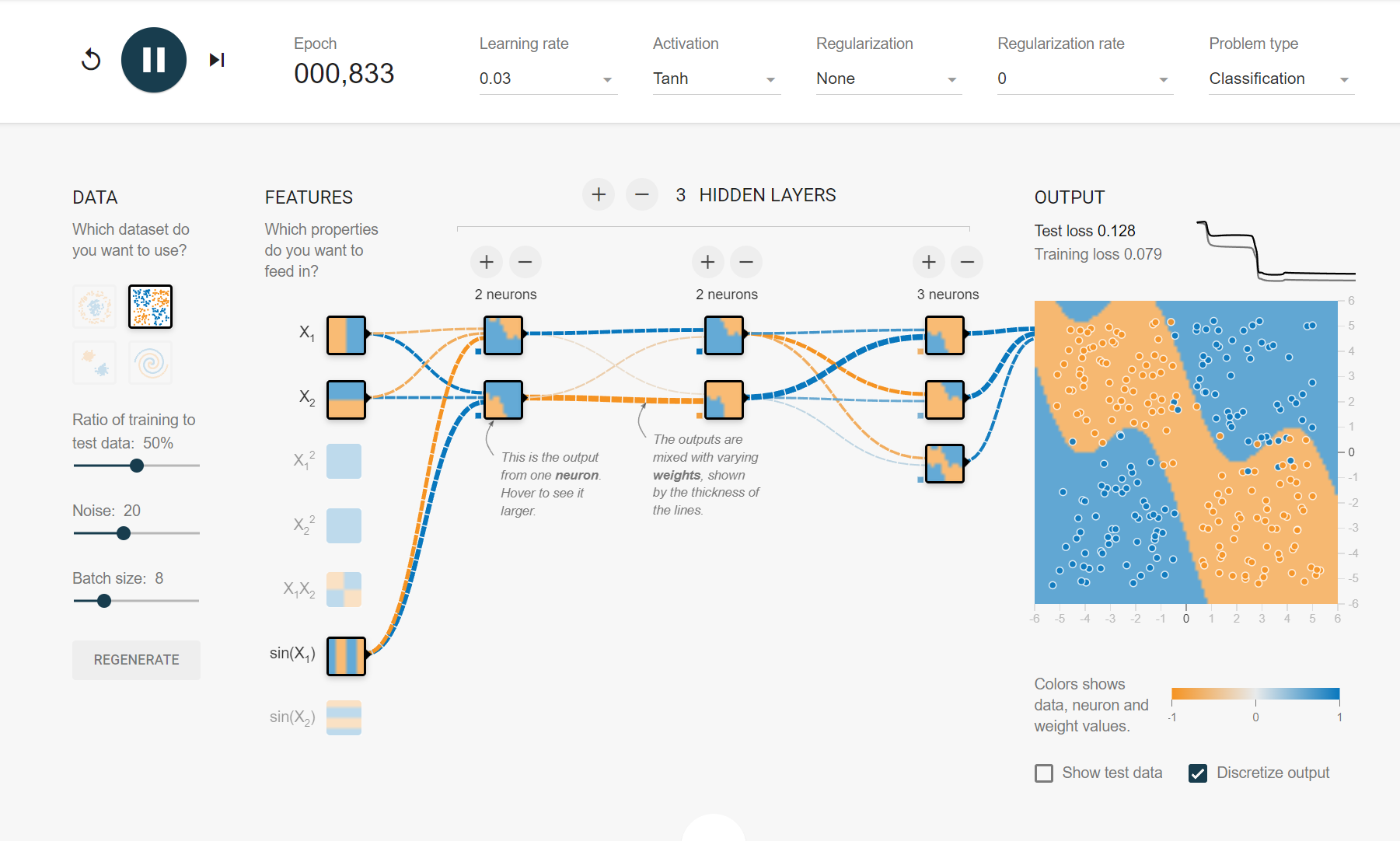


2a) Under-fitting

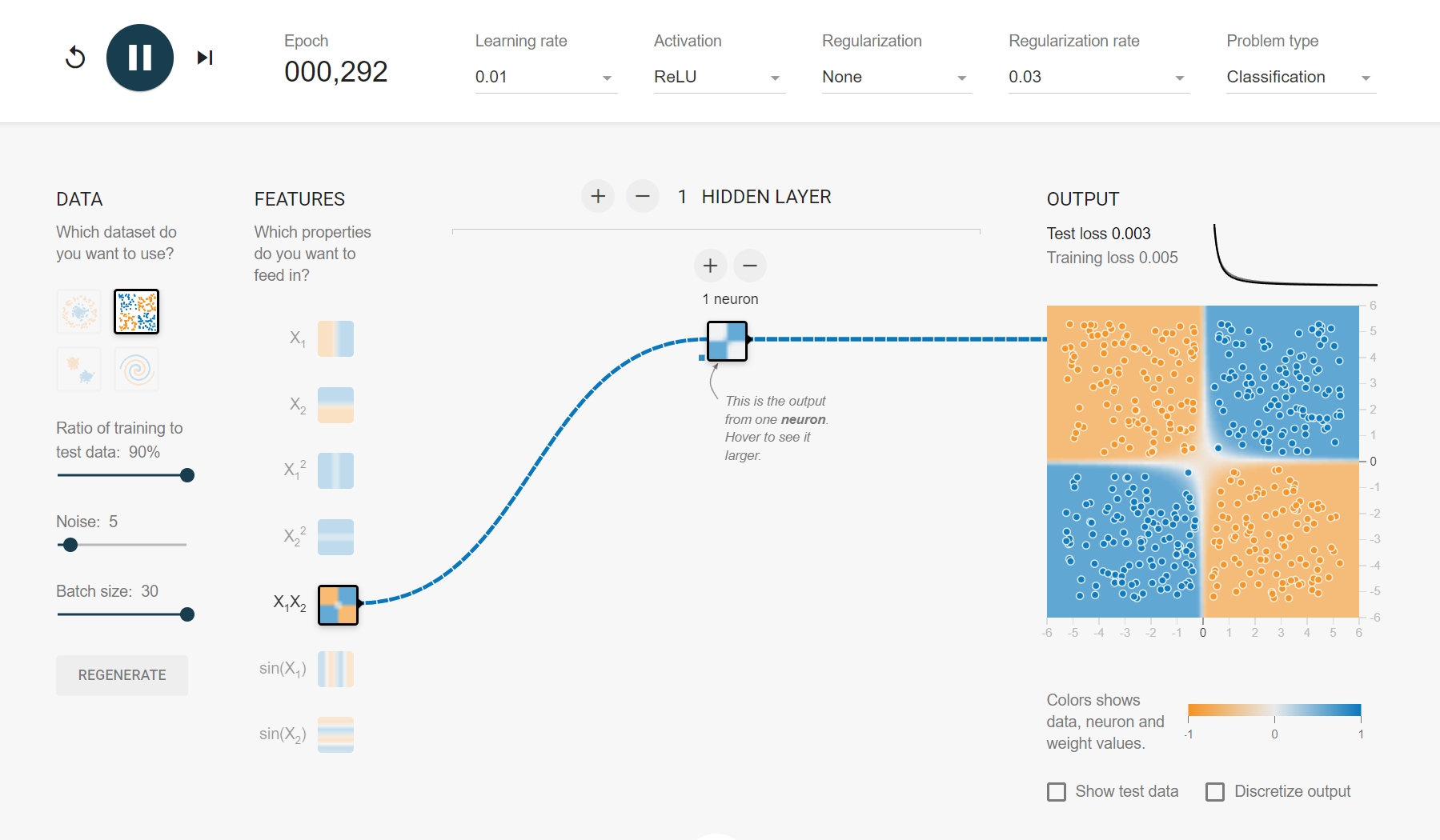
Obraz zawierający tekst, zrzut ekranu, diagram, mapa

Opis wygenerowany automatycznie

2b) Over-fitting



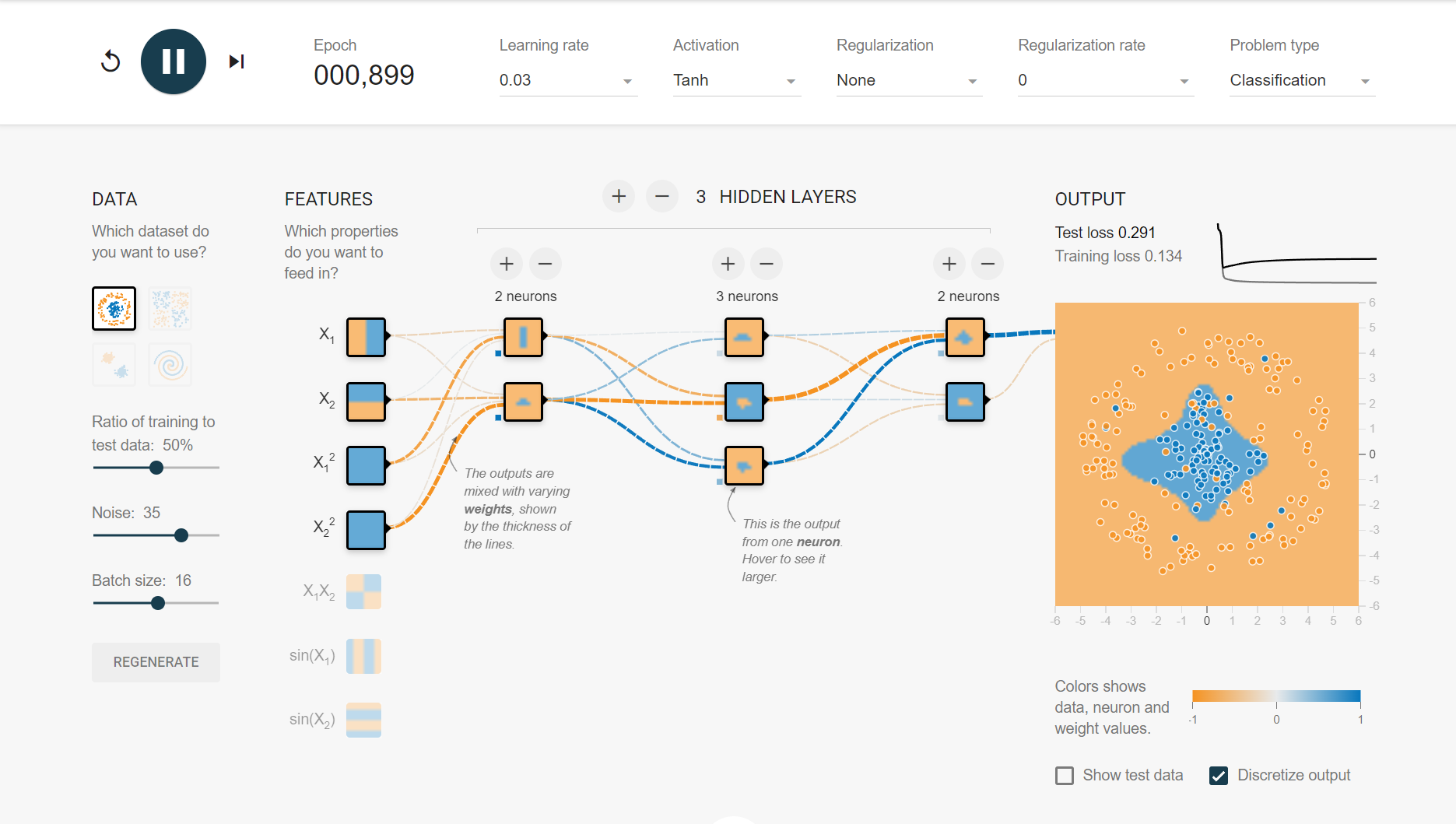
2c) Appropriate-fitting (Najlepsza minimalna architektura)



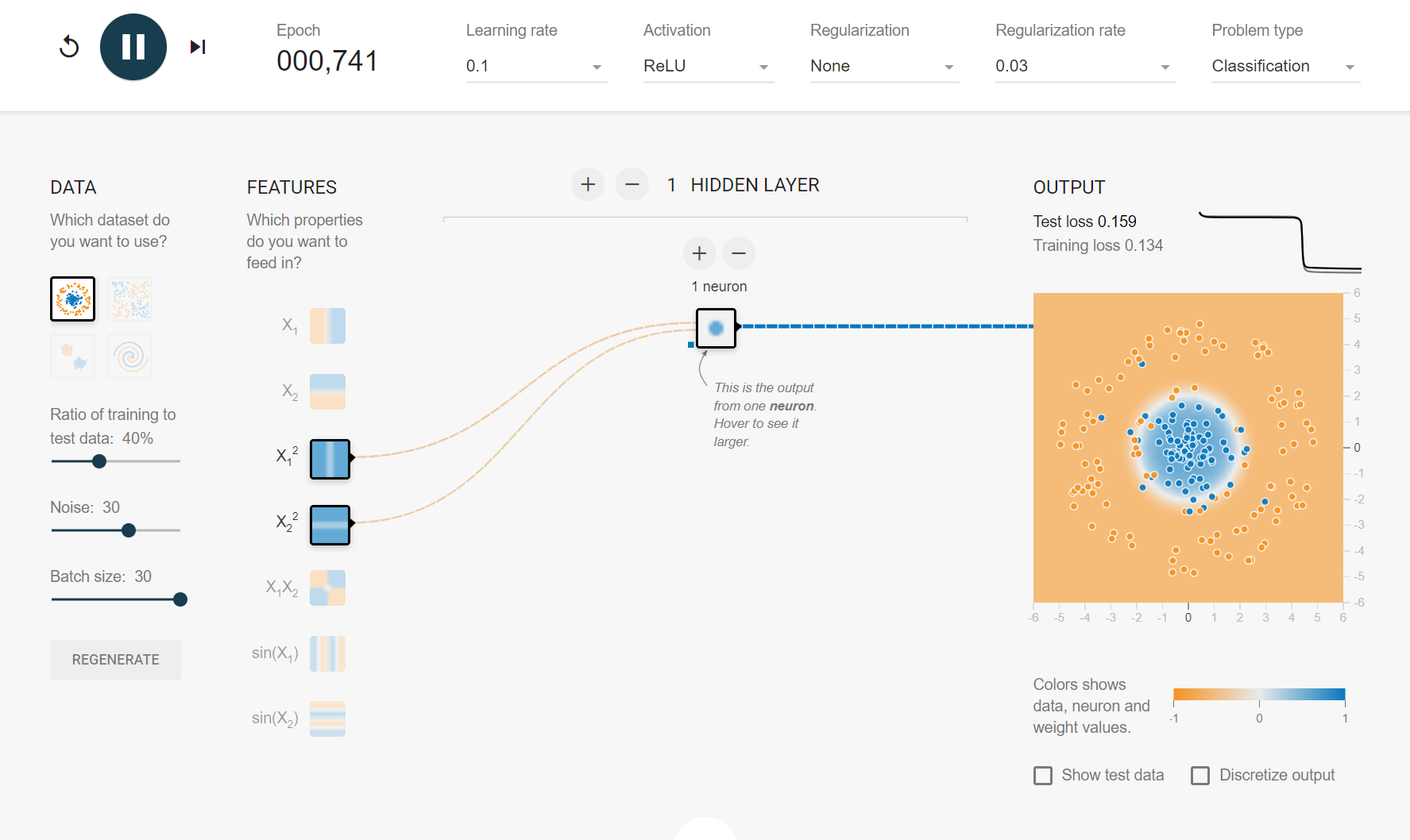
3a) Under-fitting

Niestety nie byłem w stanie stworzyć takiego modelu

3b) Over-fitting



3c) Appropriate-fitting (Najlepsza minimalna architektura)

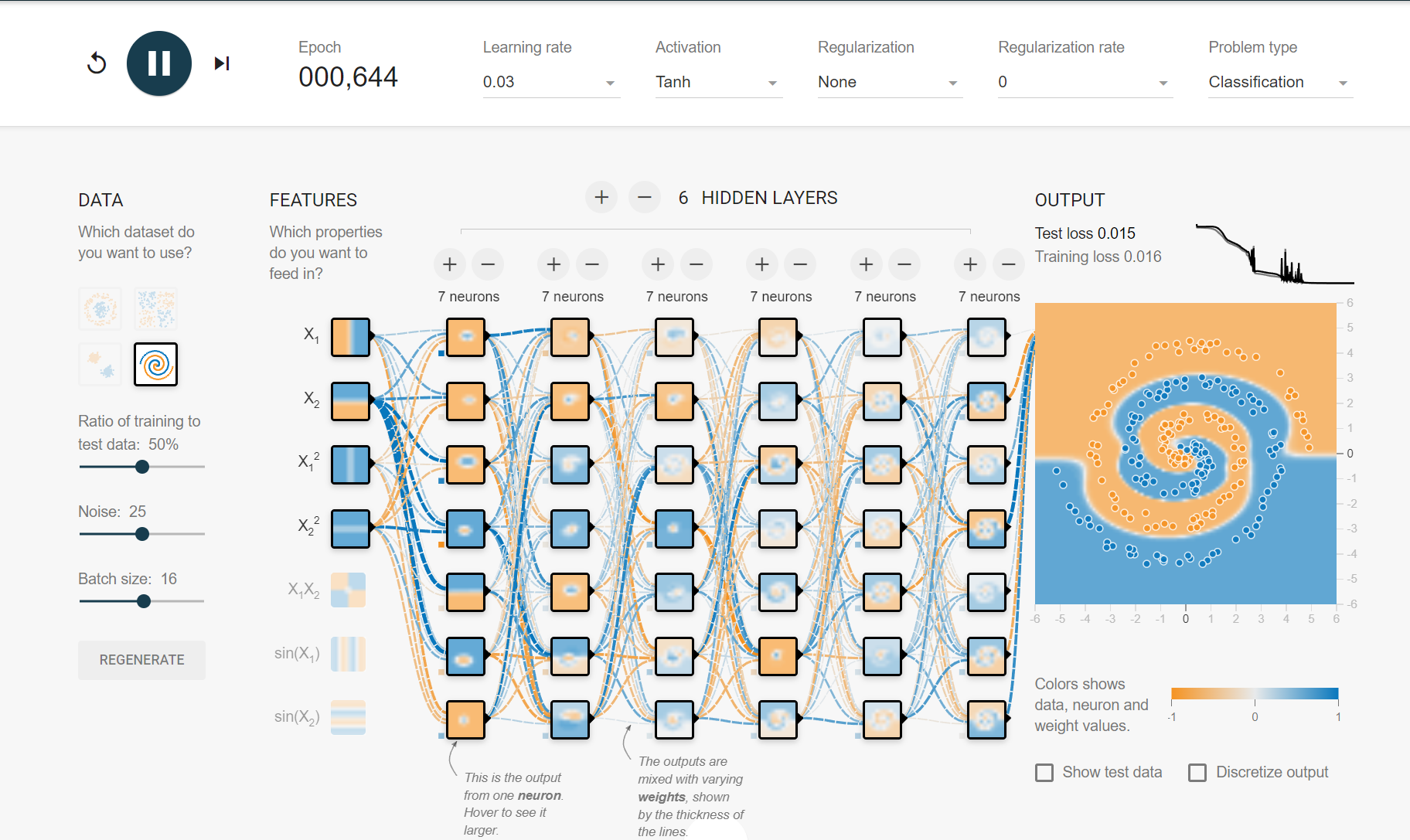


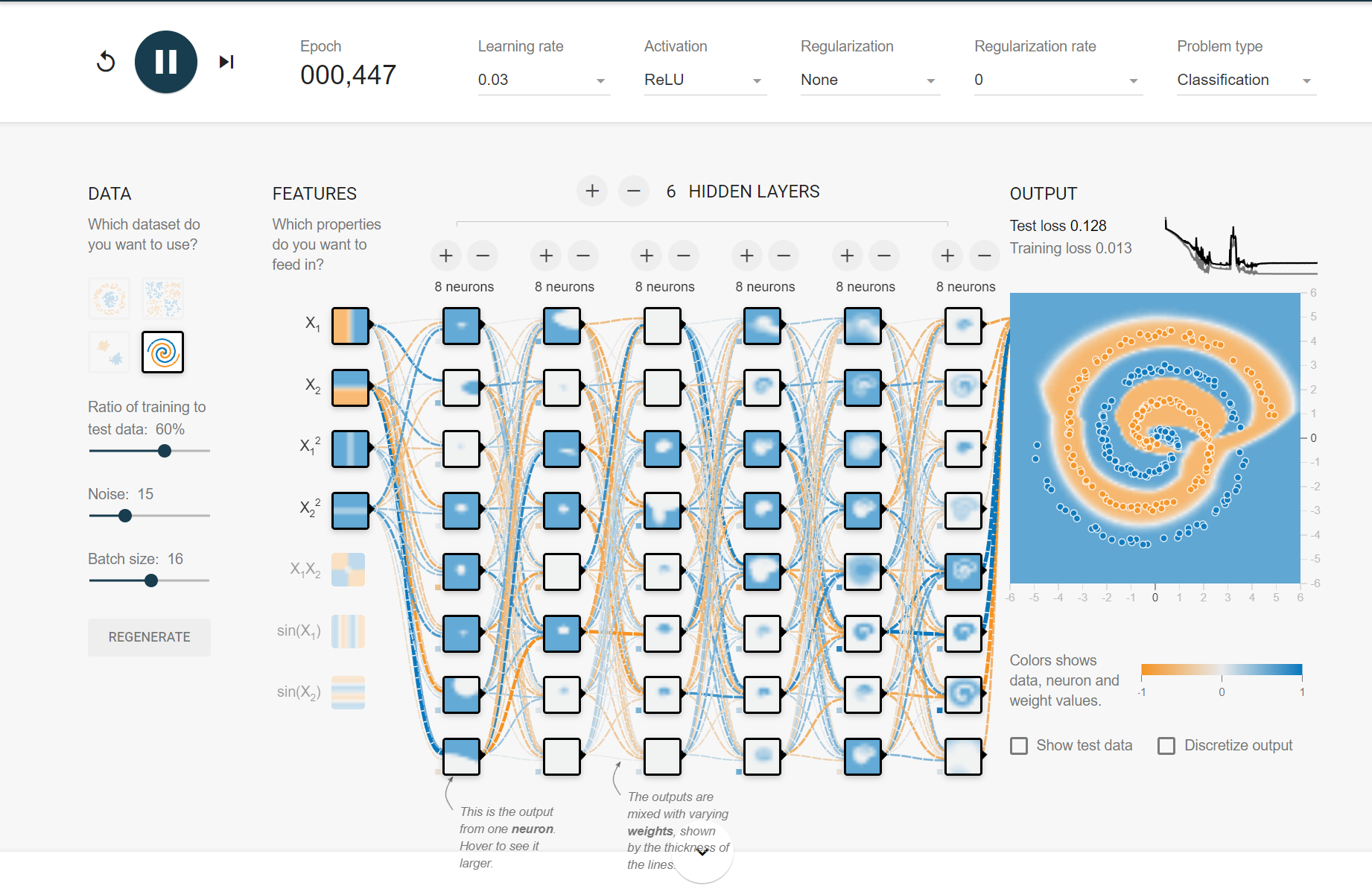
4a) Under-fitting

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

4b) Over-fitting





4c) Appropriate-fitting (Najlepsza minimalna architektura)

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

4c) Appropriate-fitting (Najlepsza minimalna architektura z drobnymi ręcznymi korektami przepustowości łączy neuronowych)

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

Wyjaśnienia pojęć oraz wskazanie za co odpowiadają poszczególne parametry modelu:

Learning rate to hiperparametr, który wpływa na tempo z którym algorytm aktualizuje wartość parametru. Learning rate wpływa na szybkość zastępowania starych danych nowymi (wpływa na prędkość uczenia się modelu).

Actiwation -  funkcja ta jest używana do wprowadzenia nieliniowości w modelu

Batch size - wpływa na rozmiar próbek

Szum - wpływa na mieszanie się próbek

Wielkość zbioru testowego - wpływa na jakość oceny modelu

Regularization to technika stosowana w uczeniu maszynowym, która polega na ograniczaniu złożoności modelu poprzez zmniejszanie wartości współczynników modelu do zera. Celem regularization jest zapobieganie nadmiernemu dopasowaniu modelu do danych treningowych, co prowadzi do gorszej generalizacji na nowych danych. [Regularization dodaje do funkcji straty modelu pewną karę, która zależy od wielkości współczynników](https://www.geeksforgeeks.org/regularization-in-machine-learning/" \t "_blank).

Istnieją trzy powszechnie stosowane techniki regularization w uczeniu maszynowym:

L1 regularization - zmusza niektóre współczynniki do bycia dokładnie zerowymi, co oznacza, że model ignoruje te cechy. [L1 regularization pomaga w selekcji cech i tworzeniu prostszych modeli](https://www.geeksforgeeks.org/regularization-in-machine-learning/)

L2 regularization - zmusza współczynniki do bycia małymi, ale nie zerowymi. [L2 regularization pomaga zmniejszyć wariancję modelu, ale zwiększa obciążenie](https://www.dataquest.io/blog/regularization-in-machine-learning/)

Elastic Net - jest to połączenie L1 i L2 regularization, które dodaje do funkcji straty kary proporcjonalne do obu norm współczynników modelu. Elastic Net pozwala na kontrolowanie stopnia zerowania i kurczenia współczynników za pomocą dodatkowego hiperparametru. [Elastic Net jest użyteczny, gdy mamy wiele skorelowanych cech lub więcej cech niż próbek](https://www.dataquest.io/blog/regularization-in-machine-learning/" \t "_blank).

[Regularization rate to hiperparametr, który kontroluje siłę wpływu kary](https://bing.com/search?q=regularization+rate+machine+learning) na model. W wyborze regularization rate istnieje kompromis między prostotą a dopasowaniem do danych treningowych:

1. Jeśli regularization rate jest zbyt wysoki, model będzie prosty, ale ryzykujemy niedopasowanie danych. [Model nie nauczy się wystarczająco dużo o danych treningowych, aby dokonywać użytecznych predykcji](https://bing.com/search?q=regularization+rate+machine+learning)
2. Jeśli regularization rate jest zbyt niski, model będzie zbyt skomplikowany i ryzykujemy nadmierne dopasowanie danych. [Model nauczy się zbyt dużo szczegółów z danych treningowych i nie będzie w stanie uogólnić na nowe dane](https://bing.com/search?q=regularization+rate+machine+learning)

Optymalna wartość regularization rate to taka, która zapewnia dobrą generalizację modelu na nowych, wcześniej niewidzianych danych.

Rozdział 3

Obraz zawierający tekst, elektronika, zrzut ekranu, oprogramowanie

Opis wygenerowany automatycznie

Kod źródłowy dla funkcji logicznej OR, NOR, AND, NAND.

Obraz zawierający tekst, zrzut ekranu

Opis wygenerowany automatycznie

Próba dwóch różnych architektur oraz zmiana liczby warstw ukrytych oraz ilość neuronów w tych warstwach

Obraz zawierający tekst, zrzut ekranu, oprogramowanie, Oprogramowanie multimedialne

Opis wygenerowany automatycznie

Obraz zawierający tekst, zrzut ekranu, Czcionka, menu

Opis wygenerowany automatycznie

1. Jaki wpływ na szybkość trenowania ma architektura sieci - Architektura sieci neuronowej znacząco wpływa na tempo trenowania modelu. Głębokie sieci wymagają więcej obliczeń, co prowadzi do dłuższego czasu trenowania. Rodzaj zastosowanych warstw, funkcji aktywacji oraz korzystanie z mechanizmów regularyzacji i transfer learningu także mają znaczenie. Wpływ na szybkość trenowania ma również rozmiar dostępnych danych treningowych. Postęp w technologii GPU i TPU przyspiesza proces trenowania, niezależnie od architektury sieci. Wybór architektury to złożony proces, uwzględniający wiele czynników, które mogą wpłynąć na tempo trenowania modelu.
2. Czy wynik końcowy jest satysfakcjonujący? - Wynik końcowy zależy od oczekiwań. Wyniki mogą być satysfakcjonujące, jeśli sieć jest w stanie dobrze dopasować się do danych treningowych.
3. Dlaczego nie można osiągnąć idealnego dopasowania wyników przewidywanych do wyników wzorcowych? - Idealne dopasowanie wyników przewidywanych do wyników wzorcowych jest niemożliwe z powodu szumu w danych, ograniczeń modelu, braku wystarczającej ilości danych treningowych, zjawiska nadmiernego dopasowania, błędów pomiarowych oraz zmienności w danych. Modele matematyczne są jedynie przybliżeniem rzeczywistych procesów, a celem jest stworzenie modelu, który skutecznie generalizuje dla nowych danych, zamiast dosłownego odzwierciedlenia danych treningowych. Unikanie nadmiernego dopasowania i radzenie sobie z różnorodnością danych są kluczowe dla skuteczności modeli.
4. Jaki wpływ na trenowanie ma wybór innego optymalizatora. - Wybór optymalizatora w trenowaniu modeli jest kluczowy. Adam przyspiesza zbieżność, lecz wymaga dokładnej kalibracji. Inne, takie jak RMSprop czy Adagrad, dostosowują tempo uczenia, ale mogą być podatne na nadmierne dopasowanie. Nadam to modyfikacja Adam, potencjalnie bardziej stabilna. Wybór zależy od specyfiki problemu, danych i architektury modelu, wymagając eksperymentów dla optymalnych wyników.

Bibliografia:

Strona internetowa z dnia 11.11.2023: <https://playground.tensorflow.org>

Strona internetowa z dnia 11.11.2023: <https://developers.google.com/machine-learning/crash-course/reducing-loss/learning-rate?hl=pl>

Strona internetowa z dnia 11.11.2023: https://developers.google.com/machine-learning/crash-course/regularization-for-simplicity/l2-regularization?hl=pl

Strona internetowa z dnia 11.11.2023: <https://www.geeksforgeeks.org/regularization-in-machine-learning/>

Strona internetowa z dnia 11.11.2023: <https://www.dataquest.io/blog/regularization-in-machine-learning/>

Sporządził:

Daniel Kasperek

Nr albumu: 129913