

Maciej Marchwiany
FUW

“Czy androidy śnią o elektronicznych owcach ?”

Problemy świadomych maszyn i ich relacje z ludźmi

Świadome maszyny, myślące roboty i podobne im byty kojarzą nam się z filmową fikcją w gatunku kina klasy B, ale czy na pewno są czystą fikcją?

Obiegowa opinia jest jednak bardziej błędna niż może się wydawać. Istnienie sztucznych sieci neuronowych, zasada ich działania, historia ich powstania i poprawność rozwiązywań otrzymywanych przy ich użyciu, pozwala już zadać pytanie o istnienie sztucznej inteligencji. Pozwala także zastanowić się nad ich świadomością, obserwowane są bowiem procesy przypisywane do tej pory tylko istotom ludzkim: marzenia.

Widać to na nawet prostych przykładach:

Rozważmy prostą sieć samoorganizującą się, która ma za zadanie rozpoznawać jaszczurkę. Początkowo sieć nie daje żadnych wyników:



Rys. 1. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci przed rozpoczęciem uczenia

Po zakończeniu uczenia sieć bez trudu rozpoznaje klasy:



Rys. 2. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci po zakończeniu uczenia

Ale jeśli zbadamy odpowiedź sieci w trakcie uczenia to często otrzymamy wynik:



Rys. 3. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci w trakcie uczenia

Sieć sama wyolbrzymiła jedna lub kilka cech jaszczurki np. wielkość ciała i kłów, sieć stworzyła smoka!

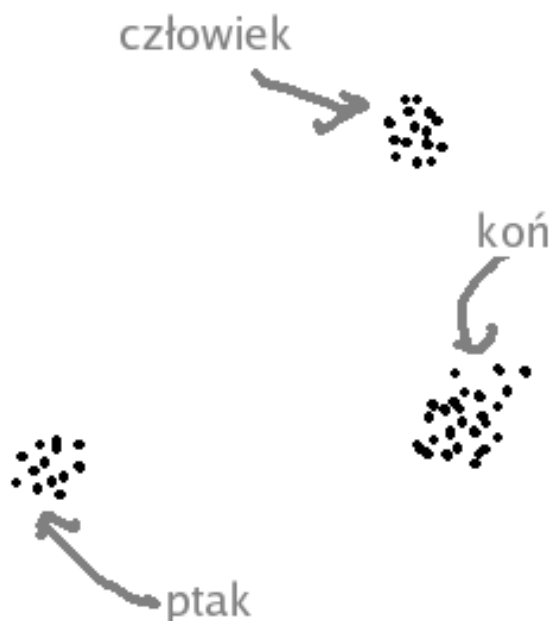
Inny przykład:

Rozważmy prostą sieć samoorganizującą się, która ma za zadanie rozpoznawać np. człowieka, konia i ptaka. Początkowo (przed uczeniem) sieć nie rozpoznaje żadnej z żądanych klas:



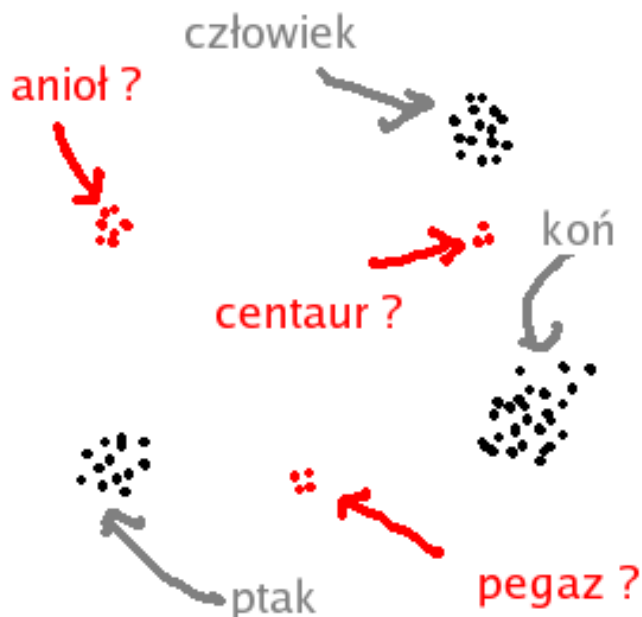
Rys. 4. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci przed rozpoczęciem uczenia

Po zakończeniu uczenia sieć bez trudu rozpoznaje klasy:



Rys. 5. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci po zakończeniu uczenia

Ale jeśli zbadamy odpowiedź sieci w trakcie uczenia to często otrzymamy wynik:



Rys. 6. Schemat przestrzeni cech z zaznaczonymi odpowiedziami sieci w trakcie uczenia

Otrzymany wynik jest dość zaskakujący, sieć rozpoznaje nieistniejące obiekty. Sieć wymyśliła coś nowego, aż ciśnie się na usta że je wymyśliła czy nawet wymarzyła. Zjawisko to jest obserwowane tylko w początkowych etapach uczenia podczas dalszej nauki efekt ten znika. Wyrażna jest tu analogia do dziecka, które marzy dorastając jednak, zdobywając coraz więcej wiedzy stopniowo przestaje.

Czy zatem marzenia maszyny są tylko fikcją? Czy nie jest już blisko do świadomości?

Uwaga:

Wszystkie rysunki są tylko poglądowymi schematami mającymi na celu zobrazowanie zjawisk