

Аналіз методів навчання на базі методів Монте-Карло для обмежених машин Больцмана

Пришляк М.Ю., *аспірант*; Субботін С.О., *професор*;

Олійник А.О., *доцент*

Запорізький національний технічний університет, м. Запоріжжя

Задачі розпізнавання та класифікації доцільно вирішувати за допомогою глибоких нейронних мереж (НМ). Усі глибокі НМ складаються з базових НМ, до яких відноситься обмежена машина Больцмана (ОМБ) – двохарова енергетична модель, яка не має зв'язків між нейронами всередині шарів, але самі шари повністю зв'язані між собою. Існуючі методи навчання ОМБ засновані на методах Монте-Карло за схемою марківських ланцюгів (МЛМК). До них відносяться: контрастивна дивергенція (КД), стійка КД, середня КД, швидка стійка КД (ШСКД) та загартовані МЛМК.

Результати проведеного аналізу показали, що:

– недолік КД полягає в тому, що зразки з негативної фази марківського ланцюга найчастіше знаходяться в безпосередній близькості до зразків з позитивної фази, що призводить до поганої працездатності вибірки Гіббса за межами навчальних даних;

– недоліком стійкої КД є нездатність використовуваного методу вибірки Гіббса в деяких випадках виходити за межі локального максимуму при багатомодальних розподілах;

– метод середньої КД має меншу похибку апроксимації, ніж КД та стійка КД, але вимагає додаткових розрахунків;

– перевагою ШСКД є постійна зміна виду розподілу в негативній фазі, але зразки на цій фазі беруться з розподілу, який може розходитися з розподілом моделі і це погіршує точність вибірки;

– основна відмінність КД, стійкої КД та ШСКД полягає в тому, як отримуються зразки на негативній фазі, але всі вони використовують однаковий механізм оновлення параметрів;

– метод загартованих МЛМК дозволяє поліпшити змішування, але витрачає більше обчислювальних ресурсів внаслідок розрахунку декількох марківських ланцюгів.

Окреслені проблеми методів навчання ОМБ обумовлюють необхідність їх поліпшення заради підвищення якості та швидкості вирішення задач розпізнавання та класифікації.