Теоремы машинного обучения

Теорема (Горбань, 1998)

Данная теорема является некоторым усовершенствованием теоремы Колмогорова. Если теорема Колмогорова по своей сути утверждает, что функцию многих переменных, определенную на единичном кубе, можно представить в виде суммы функций от одной переменной, то данная теорема идет еще дальше: она говорит, что при этом достаточно лишь линейных функций и одной нелинейной, т.е. доказывает, что неройнные сети являются эдаким универсальным аппросиматором непрерывных функций. Тем не менее, теорема ничего не гарантирует насчет размера сети.
Ссылка на оригинальную статью:
http://www.mathnet.ru/links/140e35161a69335369399cc30d8522a6/sjvm289.pdf
(http://www.mathnet.ru/links/140e35161a69335369399cc30d8522a6/sjvm289.pdf)

Теорема (Khaled, Mishchenko, Richtaric, 2019, LGD)

Данная теорема устанавливает сходимость локального градиентого спуска для функций специального вида (выпуклых и L-гладких). Она интересна тем, что сходимость устанавливается не по последней точке, как это обычно делается во всех методах спуска, а лишь по усредненной за всё время оптимизации.
Ссылка на оригинальную статью: https://arxiv.org/pdf/1909.04715.pdf (https://arxiv.org/pdf/1909.04715.pdf) (значится как **Theorem 1**)

Теорема (Каруша-Куна-Таккера)

Данная теорема устанавливает необходимые условия для точки оптимума задачи математического программирования, позволяя тем самым аналитически находить решение. Это, без сомнения, важный результат для того времени, когда методы оптимизации как наука еще только зарождались (Каруш доказал эту теорему в своей диссертации в 30-е годы, Кун и Таккер передоказали ее в 50-е годы).
Оригинальные статьи найти не удалось, вот ссылка на другую статью с доказательством:
https://folk.ntnu.no/hek/Optimering2012/kkttheoremv2012.pdf
(https://folk.ntnu.no/hek/Optimering2012/kkttheoremv2012.pdf)