Авторы выражают глубокую признательность рецензенту за его работу и полезные замечания. Все замечания с благодарностью принимаются. Соответствующие исправления внесены в работу и отмечены красным цветом. Ответы на замечания приведены ниже.

**РЕЦЕНЗИЯ**

Мне работа понравилась. Авторам удалось придумать нечто новое в достаточно популярной и в какой- то мере "заезженной" области обучения машины регрессионному и классификационному анализу. Теоретический материал заслуживает опубликования в журнале "Автоматика и телемеханика".

Однако экспериментальная часть слабовата: эксперименты проводились только для модели с одним внутренним слоем нейронов, что противоречит названию работы: где же тут глубокое, т.е. многослойное, обучение?

**Ответ:**

Выполнены эксперименты для модели с тремя внутренними слоями нейронов для выборки Boston. Результаты добавлены в Табл. 2. Эксперимент показал, что при увеличении количества слоев качество моделей увеличивается несмотря на возможность переобучения.

Отсутствуют результаты экспериментов по обучению классификации.

**Ответ:**

Результаты экспериментов по обучению классификации приведены на стр. 13 и добавлены в Табл. 2.

Изложение излишне кратко и схематично, автор недостаточное внимание уделяет повышению "читабельности" работы. Полностью отсутствует мотивационная часть. Непонятно, что нового вносит работа по сравнению с имеющимися методами;

**Ответ:**

Необходимые пояснения внесены во Введение. Новизна указана в контексте анализируемых методов.

я не нашел ответа на вопрос, возникающий у любого читателя данного журнала: зачем нужны предложенные методы, в каких классах ситуаций они позволяют давать (лучшее) решение, какого не дадут существующие методы?

**Ответ:**

Преимущества, которые дают рассматриваемые методы, описаны в конце Введения. Там же указаны классы ситуаций, в которых они дают выбор более адекватных моделей.

Имеется общее замечание по названию метода. Автор использует термин "субоптимальный" для обозначения оптимальных решений, найденных для функций, мажорируемых данной функцией. Название было бы удачным, если бы термин "субоптимальный" уже не использовался в той же области для обозначения решений, локально оптимальных для оптимизируемой функции.

**Ответ:**

Необходимые пояснения внесены после формального определения субоптимальности на стр. 5.

Далее постраничные замечания чисто редакционного порядка: В Опр. 3 - "аппроксимация" - просто замена? Если "да", зачем использовать слово "аппроксимация"? Если нет - объясните, о чем идет речь. А лучше - просто поместить опр. 3 ниже неравенства (8) и комментариев к нему.

**Ответ:**

Опр. 3 помещено ниже неравенства (8) и комментариев к нему.

Не очень понятно выражение q(w) через интеграл от q(w) по w на стр. 4 перед Утверждением 1 - порочный замкнутый круг?

**Ответ:**

Необходимые правки (шляпка q) внесены.

На стр. 5: "где расстояние DKL между двумя гауссовыми величинами задается как" лучше, чтобы было "где расстояние DKL между двумя гауссовыми величинами рассчитывается как".

**Ответ:**

Правка внесена.

После ф-лы (13) пример уместен, но очень краток. Я не понял комментарий: плохо получилось? А почему? Нужен ли метод, дающий такие плохие результаты? Может, можно дать и контр-балансирующий пример, когда метод работает хорошо?

**Ответ:**

Подробное пояснение внесено после примера. Авторы беспокоятся, что дополнительный пример необоснованно увеличит объем статьи.

То же относится к продолжению примера на стр. 8 перед секцией 3.3. А также материал непосредственно перед секцией 4 на стр. 9. Судя по всему, метод Ланжевена лучше. Зачем тогда нужен "плохой" градиентный метод?

**Ответ:**

Пояснения внесены в указанных местах. Добавлена цитата. Внесен ответ на вопрос про градиентный метод.

Стр. 7, Опр. 5: Надо ли, чтобы w-T(w) было в каком-то смысле лучше предыдущей оценки w? "следующему выражения:" - поправить на "следующему выражению:", "теорему, обобщающую технику" - неудачно сказано (теорема - утверждение, а техника - метод).

**Ответ:**

Данное требование необязательно. Все указанные правки внесены.

Стр. 8, ссылка на Ланжевина. По-моему, правильнее - да и привычнее -j писать Ланжевен (по нормам французского языка).

**Ответ:**

Фамилия исправлена по всему тексту работы.

Стр. 9 Подпись к рисунку 1 невнятна - надо четко объяснить, что иллюстрирует каждая из трех его частей, включая смысл точек, названных q(w). Вообще, идея, что подписи к рисункам должны быть понятными без необходимости изучения текста, разделяемая значительной частью международного сообщества, мне кажется и разумной, и полезной.

**Ответ**:

В подпись добавлены пояснения.

Фразу "Длина шага оптимизации у удовлетворяет следующим условиям" лучше сформулировать так "Длина шага оптимизации у удовлетворяет обычным условиям, гарантирующим сходимость метода в стандартных ситуациях", так чтобы читатель, не знакомый с этой областью, понимал, что эти условия не изобретены автором. А дальше идет что-то еще более непонятное: зачем три параметра настройки гамма? Одного не хватит? Или это тоже взято из какого-то пособия?

**Ответ:**

Исправление внесено в текст. Способ параметризации длины шага удален из текста.

Почему бы не пояснить, особенно величину тау, используемую далее? Вообще автору полезно принимать в расчет, что читатели журнала не обязательно имеют классическое математическое образование, и давать им больше пояснений.

**Ответ:**

Необходимое пояснение внесено в текст на стр. 10.