



Анализ нейросетевых моделей распознавания геомагнитных бурь

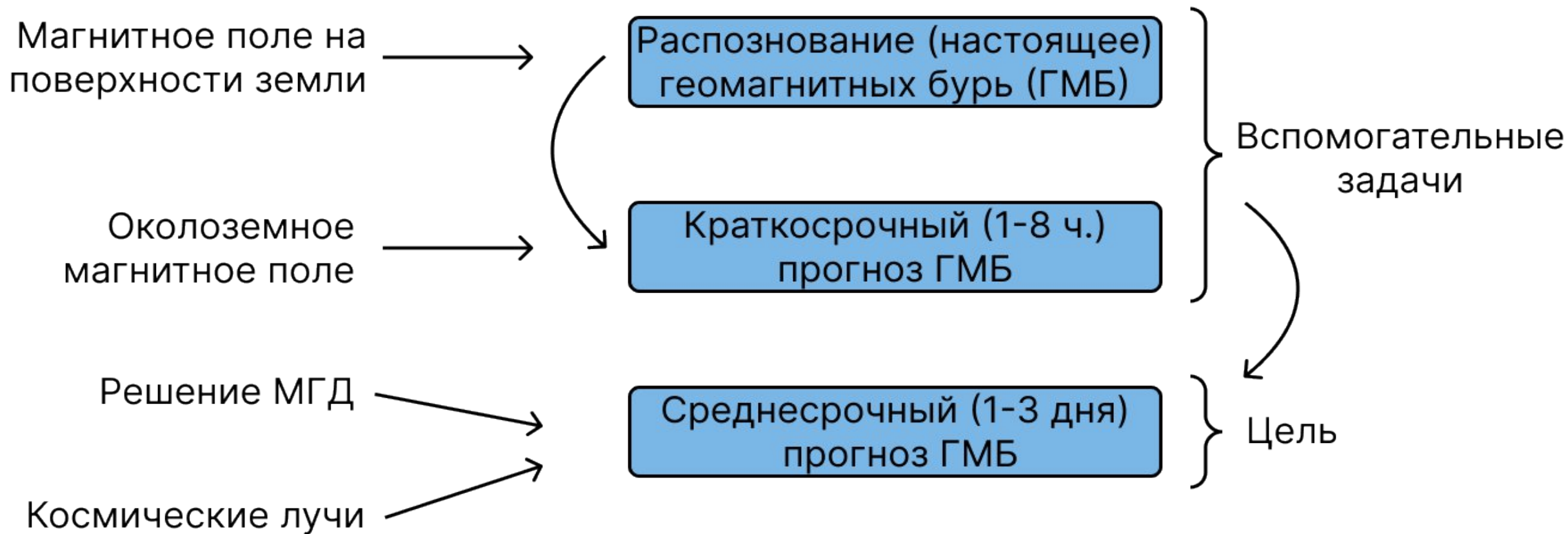
В. Чинкин^{1,2}, В. Гетманов^{1,2}, И. Яшин³, А. Ковыляева³,
А. Гвишиани^{1,2}

1. Геофизический центр РАН, Молодёжная ул 3, Москва, 119296 Россия
2. Институт физики Земли РАН, Б.Грузинская ул.10, Москва, 123242 Россия
3. Национальный исследовательский ядерный университет (МИФИ), Каширское шоссе, 31, Москва, 115409 Россия

Докладчик: Чинкин Владислав

Москва, ВККЛ-2022

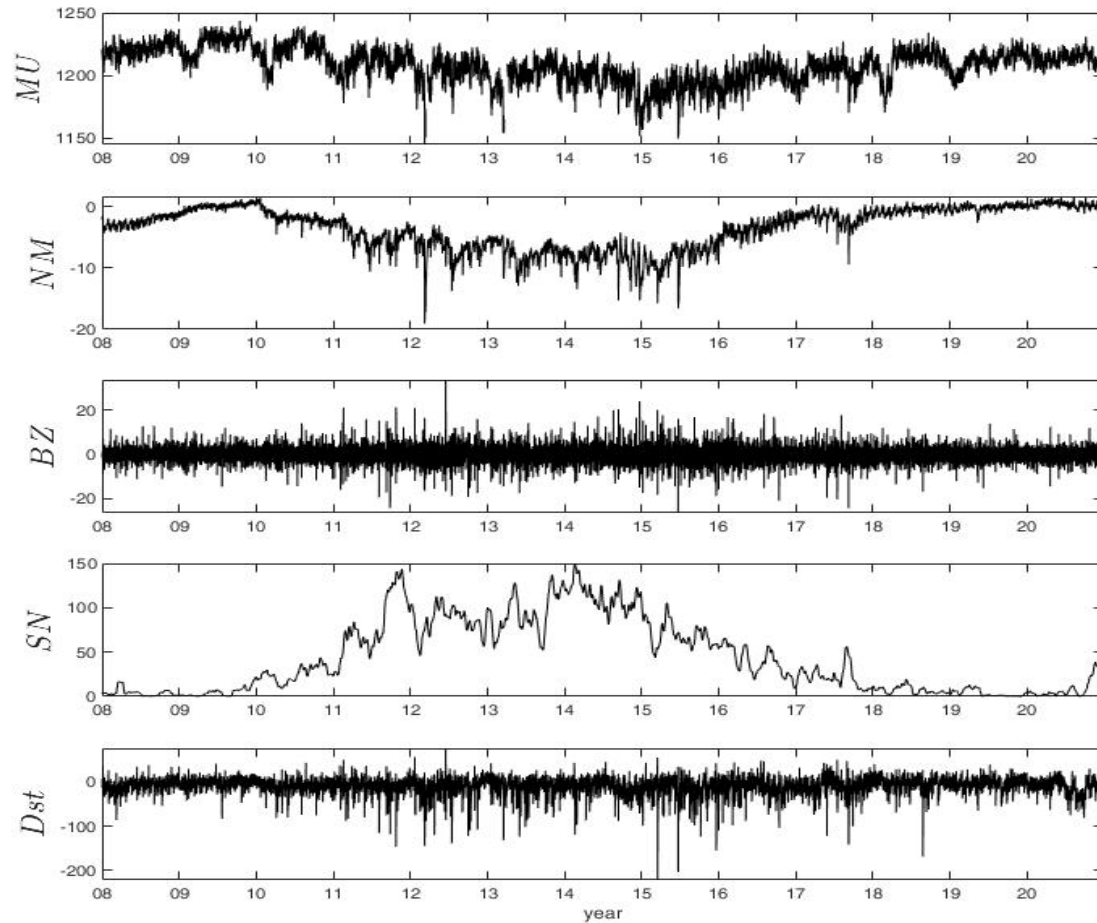
Введение



Нейросетевые модели распознавания геомагнитных бурь (ГМБ). Составляющие:

1. Входные данные (Временные ряды)
 2. Предподготовка данных:
 3. Структура нейронной сети - вид использованной нейронной сети
 4. Обучение сети
 5. Метод принятия решений распознавания ГМБ
- **Изменение одной составляющей меняет модель.**

Исходные данные

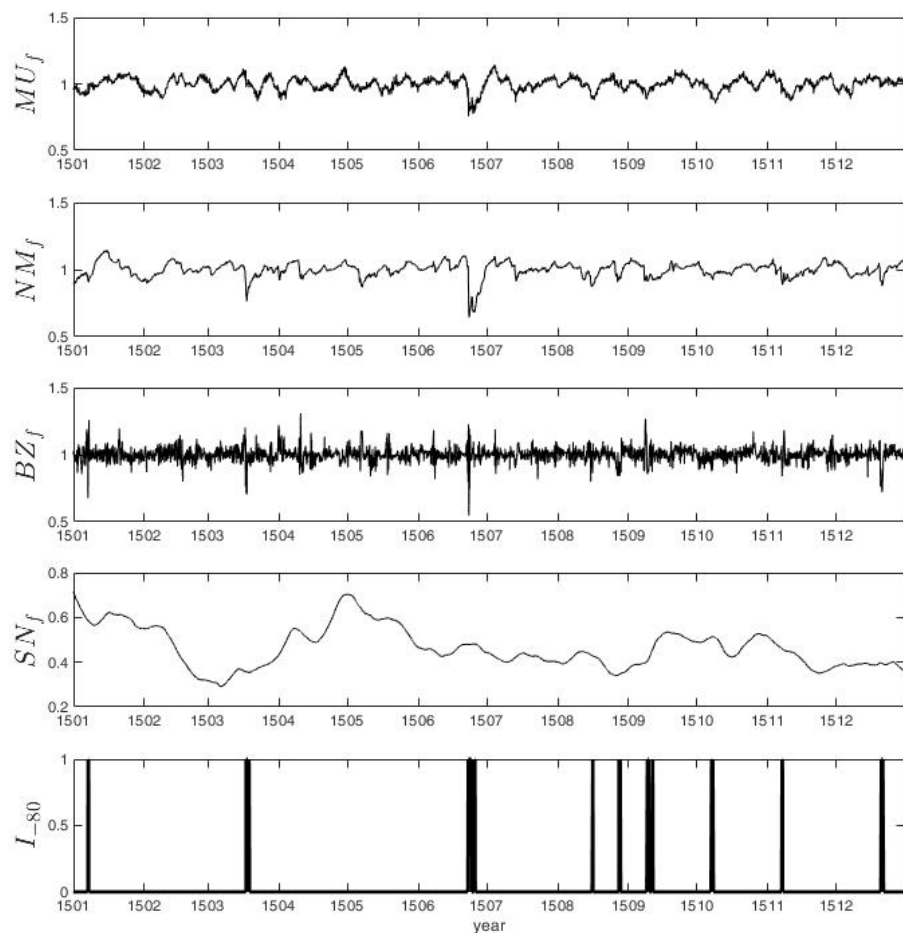


- MU(i) - Данные мюонного годоскопа УРАГАН
- NM(i) - Данные нейронных мониторов
- BZ(i) - Южная составляющая межпланетарного магнитного поля в точке Лагранжа L1
- SN(i) - сглаженные числа солнечных пятен
- Dst(i) - Dst-индекс

*Период 2008-2020гг., часовая дискретизация, i-индекс времени.

ГМБ - временной период, при котором Dst -индекс меньше определенного порога Dst_0 : $Dst < Dst_0$.

Подготовка данных



1. Заполнение пропусков
Линейная интерполяция в пропусках.
2. Фильтрация
 1. КИХ-фильтр низких частот для удаления трендов (NM, MU, BZ)
 2. КИХ-фильтр высоких частот для удаления шумов (NM, MU, BZ)
 3. Полосовой КИХ-фильтр для удаления суточной составляющей MU.
3. Нормировка
 1. Выравнивание Дисперсии
 2. Смещение данных в положительную область с центром 1 и максимальным отклонением 0.5. (NM, MU, BZ)
4. Выбор критерия магнитной бури

$$I_{Dst_0}(i) = \begin{cases} 1, & Dst(i) < Dst_0 \\ 0, & Dst(i) > Dst_0 \end{cases}$$

$I_{Dst_0}(i)$ - функция идеального распознавания по порогу Dst_0 .

Нейронные сети

Нейронная сеть - параметрическая функция, структурированная по определенным правилам.

- $f(X, W) = Y$

W – параметры (веса) сети.

Вход - X , вектор длины L .

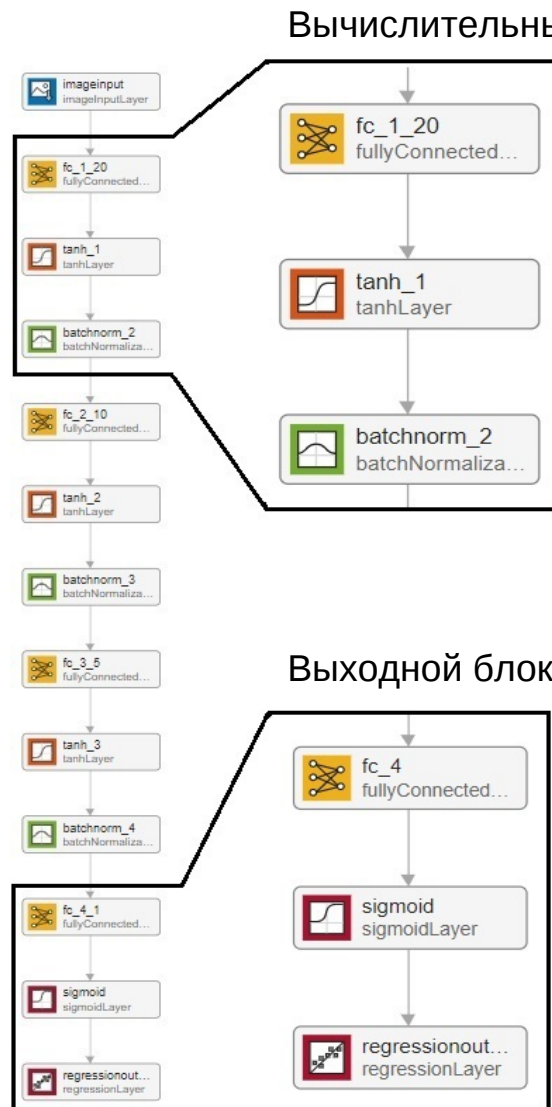
Использованные типы:

- $X_{NM}(i) = (NM_f(i), \dots, NM_f(i - L + 1)),$
- $X_{MU}(i) = (MU_f(i), \dots, MU_f(i - L + 1)),$
- $X_{NMMU}(i) = (NM_f(i), \dots, NM_f(i - L_1 + 1), MU_f(i), \dots, MU_f(i - L_2 + 1)), L_1 = L_2 = L/2.$
- и др.

Выход - $Y(i)$, число $Y(i) \in (0, 1)$

Больше $Y(i)$ – больше оценка модели, что сейчас ГМБ

Пример структуры нейронной сети



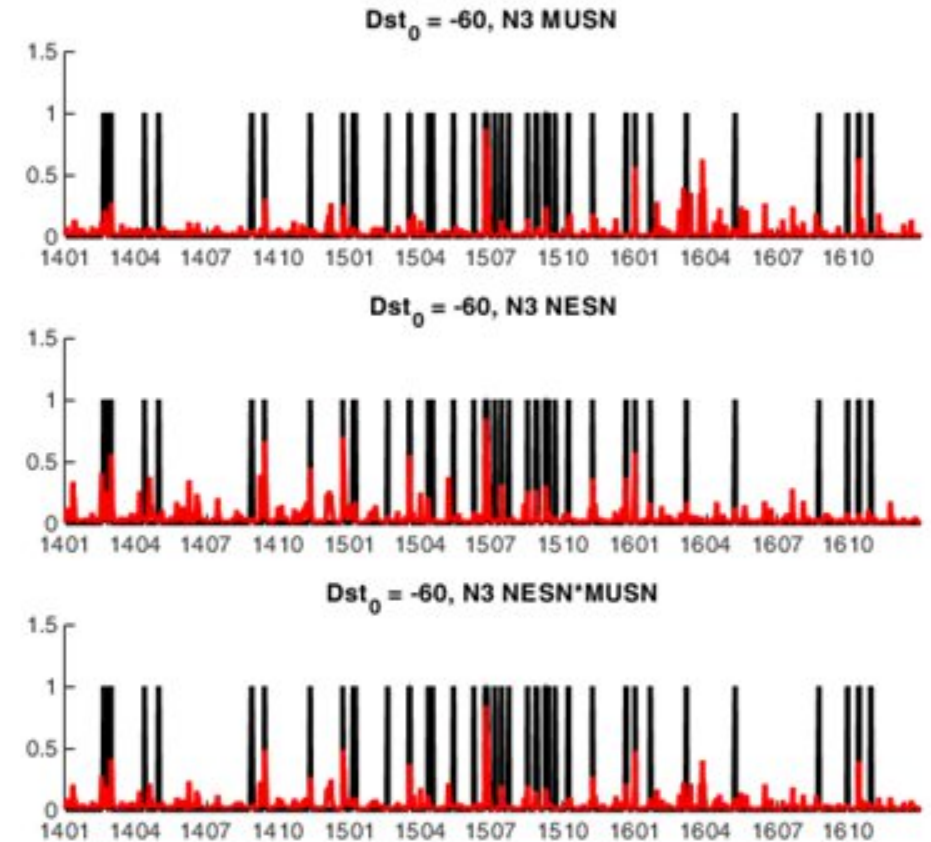
- Слой ImageInput задает размер векторов на входе.
- batchnorm_i - слои пакетной нормировки, i – номер слоя.
- fc_i_n - классические полносвязные слои нейронов, n – число нейронов в слое, i – номер слоя.
- tanh_i - функции активации тангенса гиперболического $\text{th}(x) = \frac{e^{+x} - e^{-x}}{e^{+x} + e^{-x}}$
- sigmoid, сигмоидальная функция $\text{sgm}(x) = \frac{1}{1 + e^{-x}}$
- Слой regressionoutput устанавливает конец сети
- Вычисленный блок $x_{i+1} = \text{th}(W_i x_i + b_i)$, нормировка $x'_{i+1} = \frac{x_{i+1} - \mu}{\sqrt{\sigma^2 + \epsilon}}$
 - Обучаемые параметры: W_i – матрица весов размером $n_i \times n_{i+1}$, b_i – вектор смещений размером n_i
 - Входной вектор x_i размером n_i , x_{i+1} размером n_{i+1}
 - σ , μ – параметры пакетной нормировки
- Выходной блок
 - $y = \text{sgm}(Wx_n + b_n)$

Обучение нейронной сети

- Три подтипа набора данных
 1. Обучающий набор $X(i) \in U_L: i \in I_L$
 2. Валидационный набор $X(i) \in U_V: i \in I_V$
 3. Тестовый набор $X(i) \in U_T: i \in I_T = \emptyset$
 - Весь набор данных $U = U_L \cup U_V \cup U_T$
- Обучение – процесс минимизации нормы на обучающем наборе
$$\|f(X, W) - I_{Dst0}(X)\| \xrightarrow{X \in U_1} \min$$
- Метод оптимизации – градиентный спуск с моментом
- Остановка обучения – минимальная ошибка на валидационном наборе не изменилась за N итераций

Результат обучения -

Вероятностный классификатор



Принятие решений о ГМБ-распознавании

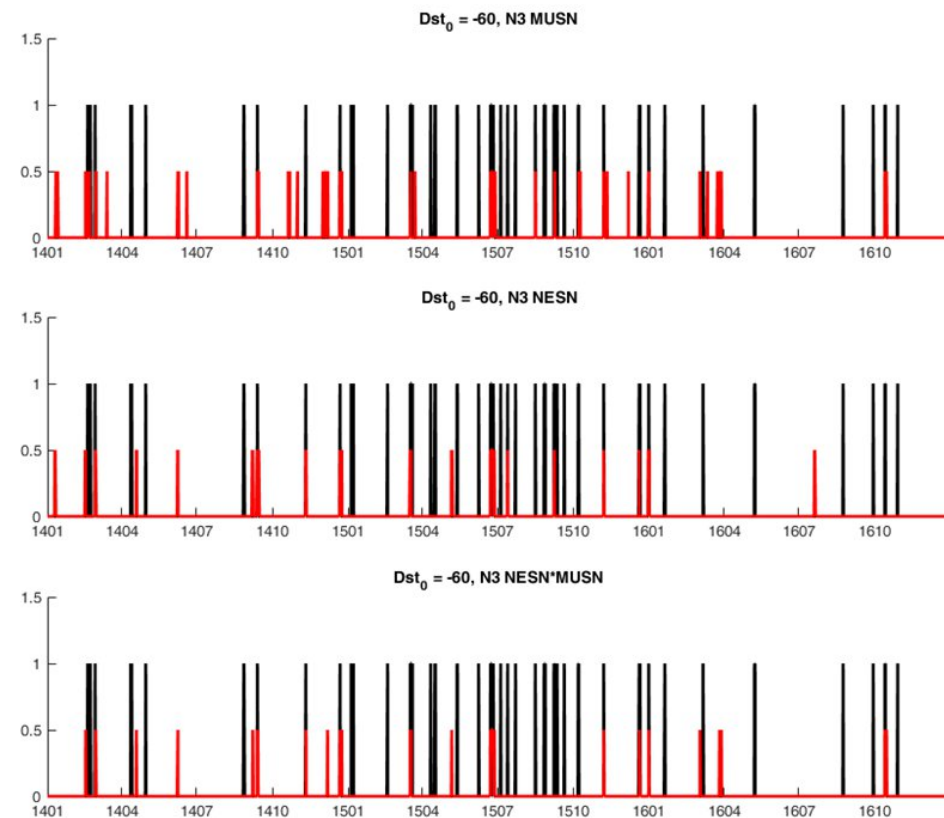
- Принятие решений – построение дискретной оценки распознавания.

- Принятие решения по порогу:

$$I^O(i) = \begin{cases} 1, & Y(i) \geq b \\ 0, & Y(i) < b \end{cases}$$

Классический пример: $b = 0.5$.

- Как выбрать наилучший порог b ?



Функционалы качества распознавания

1. Точность (Accuracy)

- $$ACC [I] = \frac{\text{Число правильных распознаваний}}{\text{Общее число распознаваний}}$$

2. True Score

- $$TS [I] = TPR [I] - FPR [I]$$

3. F1-критерий (Выбран как основной)

- $$F1 [I] = \frac{2}{\frac{1}{TPR[I]} + \frac{1}{PPV[I]}}$$

Исходные функции:

- Вероятность верного распознавания – $TPR [I]$
- Вероятность ложного распознавания – $FPR [I]$
- Ценность положительных результатов – $PPV [I] = \frac{\text{Число правильных распознаваний}}{\text{Общее число распознаваний бурь}}$

Оптимизация по параметру b для различных критериев качества

- $I^O(i) = \operatorname{argmax}_b \{ F1 [I^O(b)] \}$
- $I^O(i) = \operatorname{argmax}_b \{ TS [I^O(b)] \}$
- $I^O(i) = \operatorname{argmax}_b \{ Ac [I^O(b)] \}$

Постановка задачи анализа

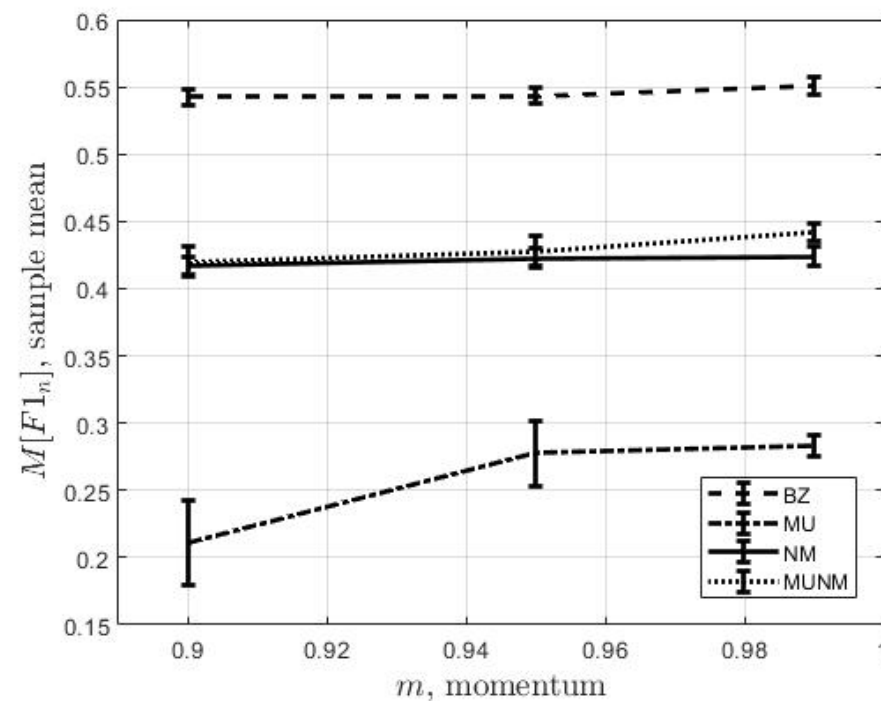
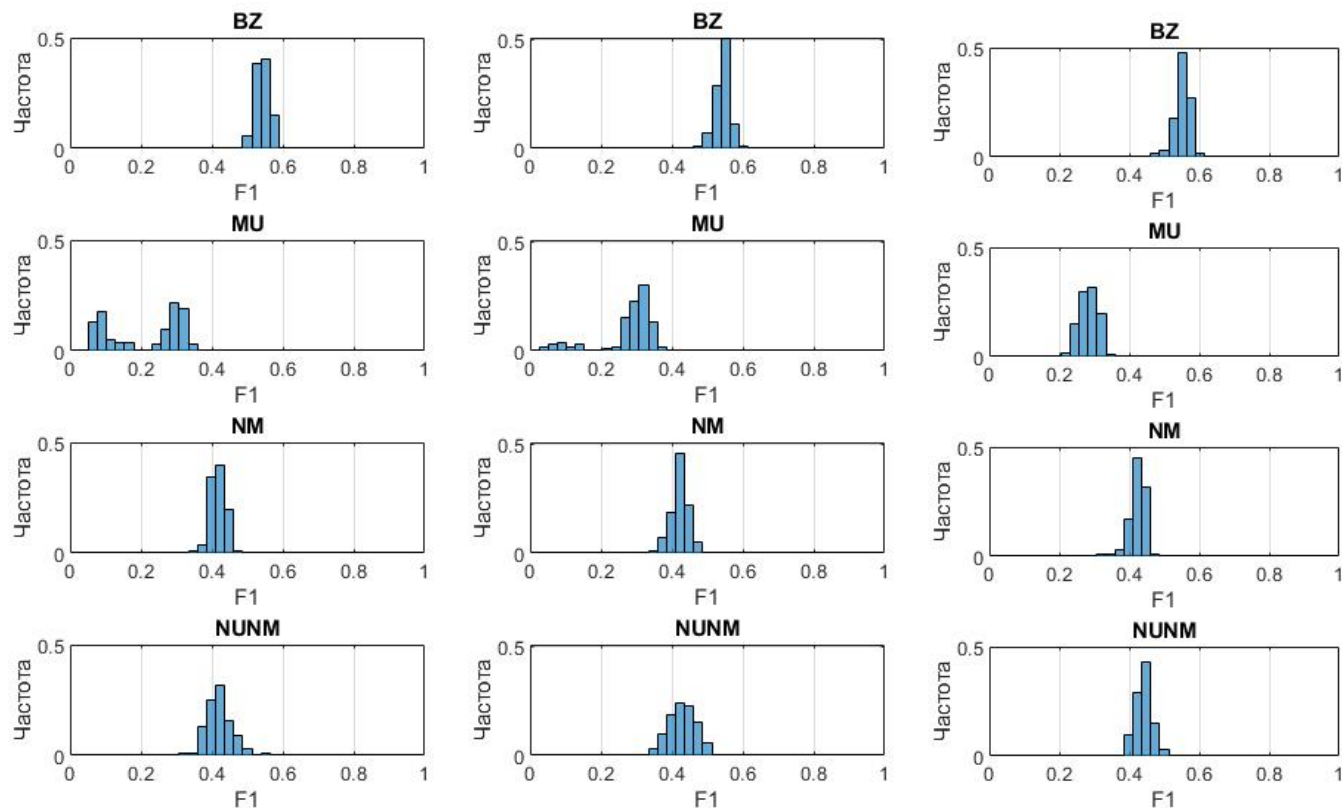
1. Построить модель распознавания ГМБ, которая максимизирует критерий качества $F1$. $I^0(i) : F1[I, I^0] \rightarrow \max$
 - $I(i)$ - функция идеального распознавания ГМБ по порогу Dst_0
 - $I^0(i)$ - оценка идентификации ГМБ
 - $q(I, I^0)$ - критерий качества классификации
2. Сравнить различные модели распознавания ГМБ по критерию $F1$ для различных элементов:
 1. Типы входных данных
 2. Структуры нейронных сетей
 3. Методы обучения

Разные типы входных данных

Момент $m = 0.9$

$m = 0.95$

$m = 0.99$



Момент m - параметр обучения, обеспечивающий выход из локальных минимумов, его увеличение повышает время обучения.

Аугментация данных: введение шума

Аугментация данных - искусственное расширение обучающего набора.

Составление дополнительных зашумленных обучающих наборов:

$$X' \in U_{L,k}: X' = X + NSR \cdot \sigma N_k(i), X \in U_L$$

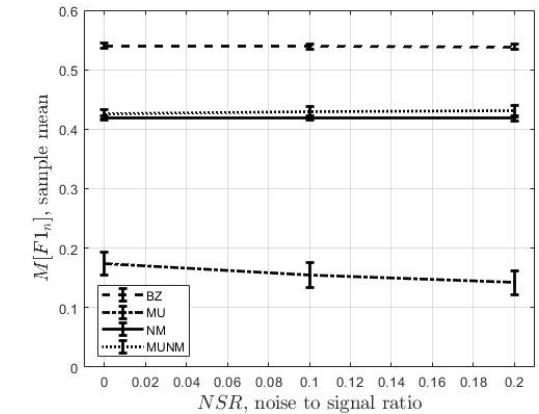
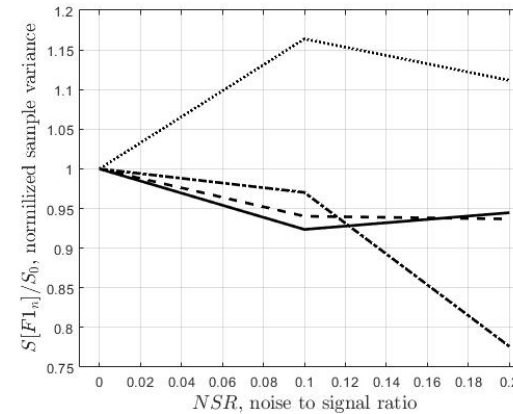
N - зашумляющий вектор, составленный из нормально распределенных случайных величин $N(0,1)$

$N(0,1)$ - нормально распределенная случайная величина с нулевым средним и единичной дисперсией.

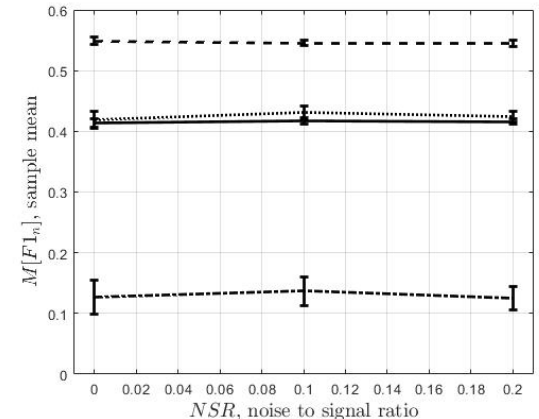
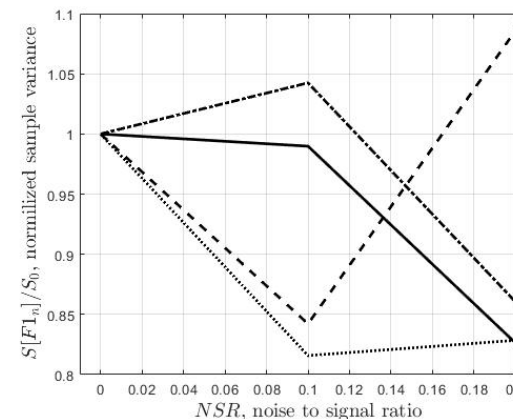
Новый обучающий набор

$$U'_L = U_L \cup U_{L,1} \cup U_{L,2} \cup U_{L,3}$$

С фильтрацией исходного шума



Без фильтрации исходного шума



Заключение

1. Анализ нейронных сетей ГМБ-распознавания позволил сделать заключение:
 1. NM-данные (нейтронных мониторов) по сравнению с MU-данными (мюонного годоскопа), обеспечили повышение эффективности ГМБ-распознавания, так как в среднем, на 30% качество распознавания, основанное на NM-данных, выше по критерию F1. Причина в том, что NM-данные были сформированы от целой глобальной сети данных нейтронных мониторов, а MU-данные в интегральном виде были сформированы от единственного мюонного годоскопа.
 2. Совместное использование NM и MU-данных дало возможность, увеличить эффективность ГМБ-распознавания, так как по критерию F1 качество распознавания, в среднем, больше на 7% по сравнению с их отдельным использованием.
 3. Эффективность ГМБ-распознавания при использовании VZ-данных (межпланетарного магнитного поля) сопоставимо с эффективностью распознавания ГМБ-распознавания на основе NM и MU –данных. Вследствие того, что VZ-данные связаны напрямую с геомагнитным полем, а NM и MU–данные связаны косвенно, качество распознавания на основе VZ-данных, выше на 20% по критерию F1.
 4. Предварительная фильтрация исходных данных положительно влияет на эффективность ГМБ-распознавания, так как увеличивает качество распознавания на 2-4% по критерию F1 для различных начальных данных.
 5. Аугментация данных при помощи введения шума положительно влияет на эффективность нефiltroванных от шума данных, увеличивая качество ГМБ-распознавания на 1-2% по критерию F1. Однако данный метод почти бесполезен или вреден для предварительно фильtroванных данных. Возможно, оптимизация данного метода позволила бы повысить качество ГМБ-распознавания нефiltroванных данных до уровня фильtroванных или выше.
2. Предложенный анализ показал, что нейросетевые модели ГМБ-распознавания имеют существенные резервы повышения эффективности на основе оптимизации параметров и являются перспективными к применению для задач прогнозирования ГМБ.

Постскрипtum

