Task 1

Пусть $x_1 < x_2 < ... < x_m, x_i \in S$. Чтобы найти ERM-гипотезу нужно посчитать ошибку на каждом из интервалов (x_i, x_j) и выбрать интервал, на котором достигается минимум. Последовательное нахождение ошибки на всех возможных интервалах выполняется за время $O(m^3)$.

Построим ДП матрицу $(l_{i,j})$, где $l_{i,j}$ - ошибка на интервале (x_i, x_j). В данном случае значения $l_{i,j}, l_{i+1,j}, l_{i,j+1}$ можно вычислить за константу. Первоначально заполняем первую строку матрицы, а затем последовательно недостающие ячейки. При заполнении каждой ячейки получившиеся значение сравниваем с текущим минимумом. Таким образом время работы алгоритма $O(m^2)$.

Task3

Пусть S размера m. Тогда в классе гипотез H_n можно выделить $C_m^n \leq m^n$ эквивалентых классов, что любые две гипотезы из одного класса одинаковы относительно S. Тогда для поиска ERM-гипотезы нужно рассмотреть по одному представителю из каждого эквивалентного класса, эмперический риск которого может быть подсчитан за $O(mn)$. Тогда ERM-гипотезу можно найти за $O(mnm^{O(n)})$.

Task4

Пусть $A \in R^{m\times n}, b \in R^m$. Пусть $b_i \neq 0$, тогда путем умножения строк матрицы на скаляр получим, что $|b_i| = |b_j| = b\ \forall i, j, \in [1, m]$.

Пусть $S = (x_i, y_i)$, где $x_i = -\text{sign}(b_i)A_i^T$, $y_i = -\text{sign}(b_i)$, A_i-i-я строка A. Покажем, что $h_{w,b}$ правильно классифицирует помеченную точку тогда и только тогда, когда i-е неравенство выполняется при w:

$$h_w(x_i) = y_i \Leftrightarrow \mathrm{y}_i(\langle w, x_i\rangle + b) > 0 \Leftrightarrow \langle w, A_i\rangle - b_i > 0$$

Таким образом, алгоритм построения ERM_{HS_n}-гипотезы для S может быть использован для решения задачи MaxFS с параметрами m, n, а также ERM_H вычислительно трудное.

Task5

Пусть $S(G)$ разделимо $(w_1, b_1), ...(w_k, b_k) \in (R^n \times R)^k$. Тогда $\forall i, y_i = 1 \Leftrightarrow \langle w_j, x_i\rangle + b_j > 0$. Рассмотрим $v_q \in V$ и поставим в соответствие e_q с меткой -1. Тогда множество $C_q = j \in [1, k] : w_j e_q + b_j < 0$ не пусто. Рассмотрим ребро $(v_q, v_r) \in E$ и поставим в соответствие пример $\frac{e_q+e_r}{2}$ с меткой $+1$. Тогда $C_q \cap C_r = \emptyset$ и можно выбрать соответствующее c_i, $\forall i \in [1, n]$, т.е. G k-раскрашиваемым.

Пусть G k-раскрашиваемый $(c_1, ..., c_n)$. К каждому вектору $w_i \in R^n$ отнесем некоторый цвет $i \in [1, k]$. Определим $w_i = \sum_{j=1}^{n} -\mathbb{I}_{[c_j = i]} e_j$. Пусть $b_i = b = 0.6$. Рассмотрим $v_i \in V$ с примером e_i. $\text{sign}(\langle w_{c_i}, e_i \rangle + b) = \text{sign}(-0.4) = -1$. Рассмотрим некоторое ребро (v_q, v_r). Так как $c_q \neq c_r$, тогда $\forall j \in [1, k]$ либо $w_{j,q} = 0$, либо $w_{j,r} = 0$. Тогда $\forall j \in [1, k]$, $\text{sign}(\langle w_j, \frac{e_q + e_r}{2} \rangle + b_j) = 1$, т.е. экземпляр связанный с ребром (v_q, v_r) классифицируется правильно. Получаем, что $S(G)$ разделимо $(w_1, b_1), ...(w_k, b_k))$. Алгоритм вычисления $ERM_{H_n^k}$ вернет ответ "граф G $k-$раскрашиваемый" тогда и только тогда, когда $S(G)$ разделимо. Учитывая то, что задача о k-раскраске является NP-трудной для $k \geq 3$, тогда задача поиска $ERM_{H_n^k}$ является NP-трудной.