Спайковые НС

Описание Spike Responce Model

Вся динамика модели нейрона *Spike Responce Model* выражается уравнением:

$$u_i(t) = u_{rest} + \sum_{j=1}^{N} w_j \sum_{f_j} exp(-\frac{t - t_j}{t_m}) a(t_j - t_i), \tag{1}$$

где потенциал покоя $u_{rest} = -70$ мВ, w_j - описывает вес j-ого синапса, t_j - время спайка на j-ом синапсе, t_i - время спайка на данном нейроне, $t_m = 20$ мс - константа описывающая угасание синаптического потенциала.

Функция $a(t_j - t_i)$ описывает феномен рефракторности конкретного синапса и отражает явления обратного распространения потенциала на синапс в случае генерирования спайка нейроном во время t_i. Такая функция описывается уравнением экспонециального восстановления:

$$a(t_j - t_i) = 1 - exp(-\frac{t_j - t_i}{t_a}), \tag{2}$$

где $t_a = 50$ мс - время восстановления синапса.

Генерация спайка производится стохастически согласно негомогенной пуассоновской плотности, описываемой через потенциал нейрона:

$$p(u) = \frac{\beta}{\alpha} ln[1 + exp(\alpha(U_{tresh} - u))] - \alpha(U_{tresh} - u) \tag{3}$$

где константы $\alpha = 1, \beta = 20$, и порог U_{tresh} по достижению которого функция резко растёт вверх равен -50 мВ.

Обучение без учителя

Обучение без учителя производится согласно оптимальной модели STDP описанной в работе Toyoizumi и др. [1]. Кратко опишем данную модель.

В качестве функции правдоподобия выбрана функция, которая отражает информацию I, которую нейрон должен максимизировать и мера выражающая отклонение частоты спайков нейрона от целевой D. Выражение D является воплощением гомеостазиса нейрона, без которого, максимизация только информации приведет к бесконечному росту синапсов. Также в формулу правдоподобия внесен фактор регулиризации распада весов т.н. *weight decay*.

Таким образом функция правдоподобия имеет вид:

$$L = I - \gamma D - \lambda \Psi \tag{4}$$

где γ - фактор влияния отклонения частоты от целевой, Ψ - регуляризационный параметр выражающий распад весов и λ - фактор, контролирующий распад.

Информацию, которую передает нейрон преобразовав входные спайки в выходные можно выразить:

$$I = \left\langle \log \frac{P(Y|X)}{P(Y)} \right\rangle_{Y,X} \tag{5}$$

где $\langle .. \rangle_X$ - среднее по ансамблю X, $P(Y|X)$ - вероятность спайков нейрона Y при данных входных спайках на синапсах X, $P(Y)$ - вероятность выходных спайков.

В рамках модели *Spike Responce Model* условную вероятность спайковой последовательности можно выразить через плотность 3:

$$P(Y_0|X) = S[0,T] = \int_0^T exp(-p(t))dt\ , \tag{6}$$

где Y_0 - нулевая спайковая последовательность или, иначе говоря, отсутствие спайков, таким образом:

$$P(Y|X) = S[0,t_{f1}]\ p(t_{f1})\ S[t_{f1},t_{f2}]\ p(t_{f2})...p(t_{fn})\ S[t_{fn},T], \tag{7}$$

что можно упростить, и считать только один интеграл:

$$P(Y|X) = S[0,T]\ p(t_{f1})\ p(t_{f2})\ ...\ p(t_{fn}) \tag{8}$$

Вероятность выходного спайка $P(Y)$ можно выразить простым средним взвешенном на определенном интервале, например 1 мин.

Гомеостатический термин можно выразить расстоянием Кульбака — Лейблера между средней частотой нейрона и целевой частотой. В качестве целевой частоты целесообразнее выбрать небольшую - до 5 Герц, в связи с оптимальностью разреженного кодирования. Таким образом выражение D имеет вид:

$$D = \left\langle \log \frac{P(Y)}{\tilde{P}} \right\rangle_{Y} \tag{9}$$

где $\tilde{P} = 5$ Герц.

Третий, регуляризационный параметр является квадратом весов синапса помноженный на термин выражающий присутствие спайка на синапсе j:

$$\Psi = \frac{1}{2}\sum_j w_j^2 \langle n_j \rangle, \tag{10}$$

здесь $n_j = 1$, если спайк присутствует на синапсе и 0 если спайк отсутствует.

Дифференцируя выражение 5 получаем:

$$\frac{dI}{dw} = \left\langle \frac{1}{P(Y|X)} \frac{dP(Y|X)}{dw_j} \log \frac{P(Y|X)}{P(Y)} \right\rangle_{Y,X} \tag{11}$$

$$\frac{dD}{dw} = \left\langle \frac{1}{P(Y|X)} \frac{dP(Y|X)}{dw_j} \log \frac{P(Y)}{\tilde{P}} \right\rangle_{Y,X} \tag{12}$$

$$\frac{d\Psi}{dw_j} = w_j \langle n_j \rangle_X \tag{13}$$

Выражения выше можно выразить при помощи специально введенных переменных $c(t)$ и $B^{post}(t)$, так что:

$$\frac{dP}{dw} = P(Y|X) \int_0^T c_j(t) dt \tag{14}$$

где

$$c_j(t) = \frac{p'_u(t)}{p(t)} [y(t) - p(t)] \int_0^\infty ds' \epsilon(s') x_j(t - s') a(s' - t_i) \tag{15}$$

а также

$$\log \frac{P(Y|X)}{P(Y)} - \gamma \, \log \frac{P(Y)}{\tilde{P}} = \int_0^T B_i^{post}(t) dt \tag{16}$$

где

$$B_i^{post}(t) = [y(t) \log \frac{p(t)}{\bar{p}(t)} - (p(t) - \tilde{p}(t))] - \gamma \, y(t) \log \frac{p(t)}{\tilde{p}(t)} \tag{17}$$

таким образом производную по весам функции правдоподобия можно выразить:

$$\frac{dL}{dw_j} = \int_0^T dt \left\langle \int_0^T c_j(t') dt' \, B_i^{post}(t) - \lambda \, w_j x_j(t) \right\rangle_{Y,X} \tag{18}$$

Дискретная симуляция модели

Получена итоговая система дифференциальных уравнений, которая пригодна для непосредственной симуляции при помощи методов интеграции:

$$\begin{cases} \frac{du_i(t)}{dt} = u_{rest} + \sum_j w_j x_j(t) a_i(t) \\ \frac{da_i(t)}{dt} = \frac{1 - a_i(t)}{t_{a_i}} \\ \frac{dx_j(t)}{dt} = -\frac{x_j(t)}{t_m} \\ \frac{dw_j(t)}{dt} = -\alpha [C_j(t) B_i^{post}(t) - \lambda \, w_j x_j(t)] \\ \frac{dC_j(t)}{dt} = -\frac{C_j(t)}{t_c} + c_j(t) \\ \frac{d\bar{p}(t)}{dt} = -\frac{\bar{p}(t)}{t_{\bar{p}}} + y_i(t) \end{cases} \tag{19}$$

здесь α - коэффициент обучения, такой что $\alpha << 1$, (был выбран 0.04), $x_j(t)$ - функция, которая равна 1, если на синапсе j в момент t есть спайк и равна нулю в ином случае. Функция $y_i(t)$ аналогичная $x_j(t)$, но описывает присутствие спайка на нейроне i. Средняя частота взвешивается на интервале $t_{\bar{p}} = 1$ мин. Взвешивание выражения $c_i(t)$ целесообразно на интервале до $t_c = 100$ мс, для более гладкой автокорреляционной функции.