



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر

پروژه تحقیقاتی درس شبکه‌های عصبی

ماشین تورینگ عصبی

نگارش

علیرضا مازوچی

استاد درس

دکتر رضا صفابخش

تیر ۱۴۰۱

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

چکیده

در این قسمت چکیده پایان نامه نوشته می‌شود. چکیده باید جامع و بیان‌کننده خلاصه‌ای از اقدامات انجام‌شده باشد. در چکیده باید از ارجاع به مرجع و ذکر روابط ریاضی، بیان تاریخچه و تعریف مسئله خودداری شود.

واژه‌های کلیدی:

کلیدواژه اول، ...، کلیدواژه پنجم (نوشتن سه تا پنج واژه کلیدی ضروری است)

صفحه	عنوان
۲	۲ ماشین تورینگ عصبی
۳	۱-۲ ساختار و محاسبات ماشین تورینگ عصبی
۵	۲-۲ وظایف یادگیری ترتیبی
۶	۳ توسعه‌های شبکه
۷	۴ کاربردها
۸	۵ پیاده‌سازی و نتایج
۹	۶ جمع‌بندی و نتیجه‌گیری
۱۰	منابع و مراجع

صفحه

فهرست اشکال

شکل

صفحه

فهرست جداول

جدول

فصل اول

مقدمه

فصل دوم

ماشین تورینگ عصبی

۱-۲ ساختار و محاسبات ماشین تورینگ عصبی

ماتع شامل دو جز است:

۱. کنترل گر شبکه که می تواند یک شبکه عصبی جلورو یا یک شبکه عصبی بازگشتی باشد.
 ۲. یک واحد حافظه خارجی که یک ماتریس حافظه $N * W$ است. N تعداد واحدهای حافظه و W ابعاد هر سلول حافظه را نمایش می دهد.
- فارغ از آنکه کنترل گر بازگشتی باشد یا خیر، کل معماری بازگشتی محسوب می شود چراکه ماتریس حافظه در طول زمان نگهداری می شود. کنترل گر سرهای خوانده و نوشتن دارد که به ماتریس حافظه دسترسی دارد. تاثیر یک عمل خواندن یا نوشتن روی یک سلول حافظه خاص با مکانیسم توجه نرم^۱ وزن دهی می شود. این مکانیسم آدرس دهی مشابه مکانیسم توجه استفاده شده در یادگیری ماشین عصبی است به جز آنکه آدرس دهی وابسته به موقعیت را با آدرس دهی وابسته به محتوای موجود در مکانیسم توجه را ترکیب می کند.
- به طور خاص برای یک ماتع در هر گام زمانی t برای هر سر خواندن و نوشتن کنترل گر یک تعدادی پارامتر را به عنوان خروجی می دهد. این پارامترها برای محاسبه وزن w_t بر روی N خانه حافظه در ماتریس حافظه M_t استفاده می شوند. نحوه محاسبه w_t در رابطه ۱-۲ آورده شده است:

$$w_t^c(i) \leftarrow \frac{\exp(\beta_t K[k_t, M_t(i)])}{\sum_{j=0}^{N-1} \exp(\beta_t K[k_t, M_t(j)])} \quad (1-2)$$

در رابطه ۱-۲ برخی از پارامترها دارای محدودیت هایی هستند:

- $\beta_t \leq 0$
- $g_t \in [0, 1]$
- $\sum_k s_t(k) = 1$
- $\forall_k s_t(k) \leq 0$
- $\gamma_t \leq 1$

در رابطه ۱-۲ w_t^c آدرس دهی وابسته به محتوا را فراهم می کند. k_t یک کلید جستجو در حافظه را نشان می دهد و K مطابق رابطه ۲-۲ یک معیار شباهت مانند شباهت کسینوسی است.

¹Soft Attention Mechanism

$$K[u, v] = \frac{uv}{||u|| \cdot ||v||} \quad (2-2)$$

با یک سری از محاسبات مطابق روابط ۲-۳، ۲-۴ و ۲-۵ ماترکها امکان تکرار بر روی وزنهای حافظه فعلی و قبلا محاسبه شده را خواهند داشت. رابطه ۲-۳ به شبکه اجازه می دهد تا بین بردار وزن قبلی یا فعلی انتخاب کند که از کدام استفاده کند. رابطه ۲-۴ امکان تکرار از طریق حافظه با عمل کانولوشن وزن فعلی و یک کرنل کانولوشنی جابجایی^۲ یک بعدی را فراهم می کند. رابطه ۲-۵ رخداد تارشدن^۳ که به واسطه عمل کانولوشن رخ داده است را اصلاح می کند.

$$w_t^g \leftarrow g_t w_t^c + (1g_t)w_{t1} \quad (3-2)$$

$$\tilde{w}_t(i) \leftarrow \sum_{j=0}^{N-1} w_t^g(j) s_t(ij) \quad (4-2)$$

$$w_t(i) \leftarrow \frac{\tilde{w}_t(i)^{\gamma_t}}{\sum_{j=0}^{N-1} \tilde{w}_t(j)^{\gamma_t}} \quad (5-2)$$

سپس بردار r_t مطابق رابطه ۲-۶ به وسیله ی یک سر^۴ خواندن خاص در زمان t محاسبه می گردد.

$$r_t \leftarrow \sum_{i=0}^{N-1} w_t(i) M_t(i) \quad (6-2)$$

نهایتا مطابق رابطه ۲-۷ و ۲-۸ هر سر نوشتن ماتریس حافظه را در گام t با محاسبه بردارهای جانبی پاک کردن یعنی e_t و جمع کردن یعنی a_t تغییر می دهد.

²Shift

³Blurring

⁴Head

$$\tilde{M}_t(i) \leftarrow M_{t-1}(i)[1 - w_t(i)e_t] \quad (۷-۲)$$

$$M_t(i) \leftarrow \tilde{M}_t(i) + w_t(i)a_t \quad (۸-۲)$$

۲-۲ وظایف یادگیری ترتیبی

برای ماتع‌ها چندین وظیفه مصنوعی در نظر گرفته شده است که تمام آن‌ها از نوع مسئله یادگیری ترتیبی^۵ است؛ زمینه‌ای که آن‌ها در آن توانمند هستند.

- رونوشت‌گیری^۶: برای وظیفه رونوشت‌گیری یک دنباله تصادفی از بردارهای بیت با یک نماد خاص به عنوان پایان دنباله به شبکه داده می‌شود. این وظیفه نیاز دارد تا شبکه دنباله ورودی را نگه دارد و سپس آن را از حافظه برگرداند.
- رونوشت‌گیری تکرارشونده^۷: مشابه وظیفه رونوشت‌گیری دنباله‌ای از بردارهای بیتی تصادفی به شبکه داده می‌شود. برخلاف وظیفه رونوشت‌گیری بعد از دنباله یک عدد که نشان دهنده تعداد دفعاتی است که باید دنباله در خروجی ظاهر شود به شبکه داده می‌شود.
- یادآوری انجمنی^۸: این وظیفه نیز یک مسئله یادگیری دنباله با دنباله‌های متشکل از بردارهای بیتی تصادفی است. در این مورد ورودی به چندین عنصر تقسیم می‌شود که هر کدام شامل بردارهای ۶×۳ بعدی است. بعد از آنکه یک دنباله از آیتم‌ها و نماد پایانی دنباله به شبکه داده می‌شود. خروجی صحیح عنصر بعدی دنباله ورودی بعد عنصر کوثری است.

^۵Sequence Learning

^۶Copy

^۷Repeat Copy

^۸Associative Recall

فصل سوم

توسعه‌های شبکه

فصل چهارم

کاربردها

فصل پنجم

پیاده‌سازی و نتایج

فصل ششم

جمع‌بندی و نتیجه‌گیری

منابع و مراجع