

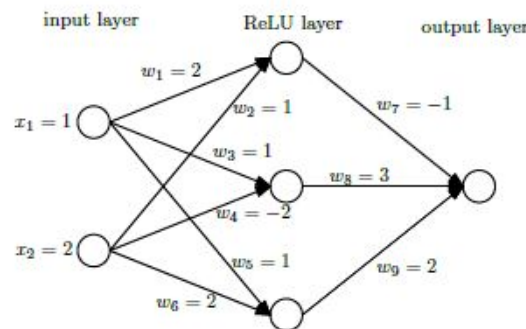
پرسش ۱: (الگوریتم پس انتشار خطا به عقب در زمان) (۴۰ نمره)
به پرسش‌های زیر پاسخ دهید.

(آ) (۵ نمره) چرا توابع فعال ساز $y = x$ و $y = \text{sign}(x)$ در لایه‌های مخفی یک شبکه چند لایه پیش‌خور مناسب نیستند؟

(ب) (۵ نمره) جمله گشتاور در الگوریتم نزول در امتداد گرادیان چه هدفی را دنبال می‌کند؟

(ج) (۱۰ نمره) چرا تابع فعال ساز Relu بیشتر از تابع فعال ساز \tanh در شبکه‌های ژرف به کار می‌رود؟ سه نمونه از برتری‌ها و کاستی‌های هر کدام را بنویسید.

(د) (۲۰ نمره) در شبکه زیر تابع هدف مربع خطا و برچسب مورد نظر برای داده ورودی ۳ است. اگر مقدار نرخ یادگیری ۰/۱ باشد با به کارگیری الگوریتم نزول در امتداد گرادیان تصادفی (SGD) مقدار تغییر وزن‌ها را محاسبه نمایید.



پرسش ۲: (Batch size) (۱۰ نمره)
در باره الگوریتم نزول در امتداد گرادیان تصادفی (SGD) به پرسش‌های زیر پاسخ دهید.

(آ) (۵ نمره) برای اندازه‌های کوچک دسته (Batch size)، چرا تعداد تکرارهای الگوریتم برای رسیدن به زیان هدف با افزایش اندازه دسته کاهش می‌یابد؟

(ب) (۵ نمره) برای اندازه‌های بزرگ دسته (Batch size)، چرا تعداد تکرارهای الگوریتم برای رسیدن به زیان هدف با افزایش اندازه دسته تغییر محسوسی نمی‌کند؟