

مصاحبه با سارا احمدی، متخصص ارشد یادگیری ماشین

تاریخ مصاحبه: ۱۵ مهر ۱۴۰۳

مدت تجربه: ۸ سال در حوزه‌های داده‌کاوی، یادگیری ماشین و یادگیری عمیق

خلاصه‌ای از مسیر شغلی

سارا احمدی دارای کارشناسی ارشد هوش مصنوعی از دانشگاه تهران است. او پس از فارغ‌التحصیلی در شرکت‌های و بینایی کامپیوتر فعالیت کرده و اکنون تیم‌های (NLP) ایرانی و سپس بین‌المللی در حوزه‌های پردازش زبان طبیعی. تحقیقاتی را در پروژه‌های بین‌المللی هدایت می‌کند.

چالش‌های کلیدی

۱. چالش‌های فنی

- کیفیت داده:** "در پروژه‌های واقعی، ۸۰٪ زمان صرف جمع‌آوری، پاک‌سازی و برچسب‌زنی داده‌ها می‌شود، نه طراحی مدل‌های پیچیده."
- توسعه مدل برای سخت‌افزار محدود:** "بسیاری از شرکت‌های ایرانی توانایی خرید سرورهای گران‌قیمت را ندارند، بنابراین باید مدل‌هایی طراحی کنیم که روی سخت‌افزار معمولی هم اجرا شوند."

۲. چالش‌های بازار کار ایران

- شکاف بین تحصیلات و صنعت:** "تئوری‌های تدریس شده در دانشگاه‌ها اغلب قدیمی‌تر از تکنولوژی‌های مورد استفاده در صنعت هستند."
- تمرکز بر توسعه پروژه‌های واقعی:** "بسیاری از شرکت‌ها به دنبال اجرای سریع پروژه هستند و فرصت تحقیق کافی نمی‌دهند."

۳. چالش‌های بین‌المللی

- رقابت سنگین:** "در موقعیت‌های بین‌المللی، با متخصصانی از سراسر جهان رقابت می‌کنید که دسترسی بهتری به داده‌ها و منابع دارند."

توصیه‌های ورود به بازار کار

۱. توسعه مهارت‌های عملی

- "پروژه‌های واقعی انجام دهید. حتی پروژه‌های کوچک اما کامل، از ده‌ها پروژه نیمه‌کاره ارزشمندترند GitHub روی."
- "مسابقه دهید، اما تنها به کسب امتیاز بالا فکر نکنید، بلکه سعی کنید راه‌حل‌های دیگران را تحلیل Kaggle در"
- "کنید."

۲. شبکه‌سازی حرفه‌ای

- "پروفایل حرفه‌ای ایجاد کنید و با متخصصان حوزه خود تعامل داشته باشید LinkedIn در"

- "در رویدادها و کنفرانس‌های آنلاین و حضوری شرکت کنید، حتی به عنوان شرکت‌کننده عادی"

یادگیری مستمر ۳.

- "هر روز حداقل یک ساعت مطالعه مقاله یا کد جدید داشته باشید"
- "و در آن عمیق (در پزشکی یا بینایی کامپیوتر در خودروهای خودران NLP مثلاً) یک تخصص خاص انتخاب کنید"
- "شوید."

بخشی از آرشیو واقعی سارا

یادداشت‌های پروژه ۱.

پروژه تشخیص احساسات از متن فارسی

تاریخ: فروردین ۱۴۰۲

چالش‌ها:

- عدم وجود دیتاست استاندارد برای احساسات فارسی
- تفاوت‌های فرهنگی در بیان احساسات
- مشکل در پردازش زبان محاوره‌ای

راه‌حل‌ها:

- ساخت دیتاست اختصاصی با ۵۰ هزار نمونه دسته‌بندی شده
- embeddings پارسی برای Bert استفاده از
- context برای درک attention اضافه کردن لایه‌های

نتایج:

- دقت نهایی: ۸۷٪ (۱۰٪ بهتر از مدل‌های قبلی)
- زمان پردازش: ۰.۲ ثانیه برای هر متن

برنامه روزانه یادگیری (نمونه یک هفته) ۲.

منبع	زمان	موضوع	روز
arXiv مقاله از	۲ ساعت	مدل‌های Fine-tuning Vision Transformer	شنبه
کد شخصی	۳ ساعت	کار روی پروژه شخصی - سیستم توصیه‌گر	یکشنبه
وبلاگ‌های تخصصی	۱.۵ ساعت	مطالعه معماری‌های جدید NLP	دوشنبه
رویداد آنلاین	۲ ساعت	شرکت در وبینار بین‌المللی	سه‌شنبه
کد شرکت	۲ ساعت	رفع باگ پروژه کاری	چهارشنبه

کد نمونه از یک پروژه واقعی ۳.

پیش‌پردازش داده برای پروژه تشخیص تصادفات ترافیکی

```
def preprocess_traffic_images(image_paths, labels):
```

```
    """
```

```
    پیش‌پردازش تصاویر ترافیکی برای مدل تشخیص تصادف
```

```
    """
```

```
    processed_images = []
```

```
    processed_labels = []
```

```
    for img_path, label in zip(image_paths, labels):
```

```
        # خواندن تصویر
```

```
        img = cv2.imread(img_path)
```

```
        # افزایش کنتراست برای شرایط نوری مختلف
```

```
        img = enhance_contrast(img)
```

```
        # (ROI) تشخیص خودکار منطقه مورد علاقه
```

```
        roi = detect_roi(img)
```

```
        # نرمال‌سازی
```

```
        img_normalized = normalize_pixels(roi)
```

```
    processed_images.append(img_normalized)
```

```
    processed_labels.append(label)
```

```
return np.array(processed_images), np.array(processed_labels)
```

لیست منابع پیشنهادی ۴.

- **دوره‌های آنلاین:**
 - Coursera: Deep Learning Specialization توسط Andrew Ng
 - Fast.ai: Practical Deep Learning for Coders
- **کتاب‌ها:**
 - "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow"
 - "Deep Learning for Computer Vision"
- **کانال‌های ایرانی:**
 - هوش مصنوعی به زبان فارسی (تلگرام)
 - ایران‌ای‌آی (وبسایت)

نکته پایانی از سارا

ورود به بازار کار هوش مصنوعی نیازمند صبر و پشتکار است. شکست‌های اولیه بخشی از مسیر یادگیری هستند. مهم این است که هر پروژه، حتی اگر ناموفق باشد، به دانش و تجربه شما اضافه می‌کند. حوزه هوش مصنوعی به سرعت در حال تغییر است، بنابراین اگر عاشق یادگیری مستمر نیستید، این حوزه برای شما مناسب نیست."
