



به نام خدا

گروه مهندسی برق و کامپیوتر

دانشکده فنی و مهندسی

عنوان پروژه کارشناسی: تخصیص دریکله پنهان (LDA)

چکیده:

تخصیص دریکله پنهان^۱ (LDA) روشی برای تعیین موضوع یک سند^۲ است که بر اساس آمار و احتمالات بیشترین ارتباط معنایی یک متن با یک موضوع را مشخص می کند. ابتدا فایل ورودی (سند) خوانده و کلمات^۳ موجود در سند به طور تصادفی به موضوعاتی^۴ که از قبل تعیین شده اختصاص داده می شوند سپس بر اساس یک تخصیص دریکله و احتمالاتی که ماهیت LDA را تشکیل می دهند و با چندبار تکرار شدن یک حلقه شرطی، هر کلمه به موضوعی اختصاص داده می شود که بیشترین ارتباط را با آن دارد، در آخر میزان ارتباط معنایی سند ورودی با هر یک از موضوعات توسط اعداد و به صورت کمی بیان می شود

استاد راهنما: دکتر میرمحسن پدram

ارائه دهنده: ندا طیبی، عاطفه ثقفی

۱. Latent Dirichlet Allocation

۲. Document

۳. Word

۴. Topic