تشخیص موضوع یکی از مسائل حوزه-ی پردازش زبان طبیعی است که در سال-های اخیر مورد توجه قرار گرفته و پژوهش-های گوناگونی بر روی آن انجام شده است. در این مسئله هدف اصلی خوشه-بندی اسناد متنی در دسته-های مختلف است به نحوی که اسناد هر خوشه از نظر موضوعی به هم نزدیک باشند. راه حل-های مختلفی برای این مسئله ارائه شده است که بخش قابل توجهی از آن-ها از الگوریتم-های خوشه-بندی مانند K-means برای حل این مسئله استفاده می-کنند. علاوه بر روش-های مبتنی بر خوشه-بندی اسناد در برخی از پژوهش-ها از روش-های مدل-سازی موضوعی برای حل این مسئله استفاده شده است. در این پژوهش دو روش تشخیص موضوع ارائه می-شود که مبتنی بر الگوریتم خوشه-بندی K-means عمل می-کنند و با ارائه-ی روش-هایی برای تعیین مراکز اولیه مناسب برای این الگوریتم کیفیت آن را بهبود می-دهند. با توجه به حساسیت قابل توجه الگوریتم K-means به انتخاب مراکز اولیه، در این پژوهش ابتدا این میزان از حساسیت را بصورت عملی در مسئله تشخیص موضوع نشان می-دهیم و سپس دو روش برای انتخاب هوشمندانه مراکز اولیه در الگوریتم K-means ارائه می-شود که باعث بهبود کیفیت در مسئله-ی تشخیص موضوع می-شوند. در روش پیشنهادی اول برای انتخاب مراکز اولیه ابتدا گراف اخبار را ساخته و بر روی آن الگوریتم DivRank را که یک روش گام-برداری تصادفی تقویتی است اجرا می-کنیم تا با استفاده از آن مراکز اولیه مناسب را شناسایی کنیم. در روش پیشنهادی دوم نیز یک الگوریتم ارائه شده است که با بهره-گیری از مدل-سازی موضوعی (LDA (Latent Dirichlet Allocation اقدام به انتخاب هوشمندانه-ی مراکز اولیه می-کند و در نهایت با داشتن مراکز اولیه مناسب اسناد را خوشه-بندی می-کنیم. در روش-های ارائه شده برای محاسبه فاصله اسناد از توزیع موضوع حاصل از LDA آن-ها استفاده شده است. در نهایت به انجام آزمایش بر روی سه مجموعه دادگان مختلف پرداختیم، آزمایش-ها نشان می-دهند که استفاده از روش-های ارائه شده در این پژوهش باعث بهبود چشم-گیر نتایج نسبت به روش LDAدر دو مجموعه از سه مجموعه دادگان می-شود. همچنین نتایج نشان می-دهند که مراکز اولیه حاصل از روش-های پیشنهادی برای انتخاب مراکز اولیه در مقایسه با روش-های انتخاب مراکز اولیه تصادفی و روش K-mean++ در دو مجموعه دادگان همیشه مناسب-تر بوده و احتمال بهتر بودن مراکز اولیه انتخابی در مجموعه دادگان دیگر مورد آزمایش بیش از ۷۰ درصد است.