Outfit recommendation is a task of suggesting fashion products that complement a given set of garments. Traditional recommender systems rely primarily on similarities between items or users; however, that is not sufficient for a recommendation of complementary products. Thus, outfit recommendation systems use machine learning techniques to learn more subtle relations between items. In this thesis, we explore the possibility of employing recent natural language processing approaches in outfit recommendation. We propose a novel architecture based on the Transformer, and we evaluate the model on standard datasets. We show that our approach is capable of learning some relations between items. However, its performance does not exceed the state-of-the-art models.