يعد نظام كشف التسلل أحد خطوط الدفاع الرئيسية لتوفير الأمن للبيانات والمعلومات وشبكات الكمبيوتر. تتمثل إحدى مشكلات أنظمة الأمان هذه في زيادة وقت المعالجة وانخفاض معدل الكشف الذي يحدث بسبب الكمية الكبيرة من البيانات التي تحتوي على العديد من الميزات غير الملائمة والمتكررة. لذلك ، يمكن أن يؤدي اختيار الميزة إلى حل هذه المشكلة عن طريق تقليل عدد الميزات. يعد اختيار طرق اختيار الميزة المناسبة التي يمكن أن تقلل من عدد الميزات دون التأثير السلبي على دقة التصنيف تحديًا كبيرًا. حفزنا هذا التحدي على التحقيق في تطبيق تقنيات اختيار الميزات المختلفة في اكتشاف التسلل. قمنا بتحليل أداء التقنيات المختارة مثل البحث الجشع والقضاء الخلفي والخوارزمية الجينية ومعالجتها ومقارنتها بالتقنيات الحالية. تم استخدام عدة قواعد بيانات لاجراء التجارب والتحقق من النتائج. أظهرت النتائج فعالية الخوارزمية الجينية وتفوقها على الخوارزميات الأخرى المستخدمة في الدراسة. حيث حصلت الخوارزمية الجينية على معدل دقة أعلى ومعدل خطأ أقل من الخوارزميات الاخرى في أكثر من قاعدة بيانات مستخدمة في الدراسة.

An intrusion detection system (IDS) is one of the main defense lines that provide security to data, information, and computer networks. The problems of this security system are increased processing time, high false alarm rate, and low detection rate that are caused by a large amount of data that contain various irrelevant and redundant features. Feature selection (FS) can solve these problems by reducing the number of features. Choosing appropriate feature selection methods that can reduce the number of features without a negative effect on classification accuracy is a major challenge. This challenge motivated us to investigate the application of different FS techniques in intrusion detection. The performances of the selected techniques such as genetic algorithm, greedy search, and back elimination are analyzed, addressed, and compared with existing techniques. Several machine learning methods such as support vector machine (SVM) and multilayer perceptron (MLP) are used as objective function for selected FS techniques as well as classification process. The CIC-IDS-2017, CSE-CIC-IDS-218, and NSL-KDD datasets are considered for the experiments. The efficiency of the proposed models was proved in the experimental results, which indicated that it had highest accuracy in the selected detests